Unidirectional CI and the consequences of *Wolbachia* for gene flow and reinforcement

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

von
Dipl.-Biophys. Matthias Flor
geboren am 23.11.1973 in Kiel

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. h.c. Christoph Markschies

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. Lutz-Helmut Schön

Gutachter:
1. Prof. Dr. Peter Hammerstein
2. Prof. Dr. John H. Werren
3. Prof. Dr. Dr. h.c. Edda Klipp

eingereicht am: 4. Februar 2010
Tag der mündlichen Prüfung: 29. März 2011

Abstract

The intracellular bacterial parasites of the genus *Wolbachia* are widespread among arthropod species. In many hosts, they induce a reproductive incompatibility between uninfected females and infected males. The potential role of this cytoplasmic incompatibility in speciation processes of the bacteria's hosts has long been debated. In this thesis, we analyze common criticisms of such a role by means of mathematical models, combining *Wolbachia* infection dynamics and host population genetics.

In particular, we are concerned with the following: (i) In order to measure the stability of infection patterns within host metapopulations, we derive critical migration rates. (ii) We evaluate the impact of cytoplasmic incompatibility on gene flow between populations by calculating effective migration rates. (iii) We determine the conditions that favor the evolution of female mating preferences through reinforcement. Finally, (iv) we apply our models to a particular real-world speciation process of two sibling *Drosophila* species in North America, discuss emerging problems, and suggest future directions of research.

In summary, our results implicate that *Wolbachia* might be a frequent factor in host speciation, but usually only by contributing to overall reproductive isolation among other factors. Reinforcement of premating isolation is selected for only under stringent conditions.

Keywords:
Wolbachia, cytoplasmic incompatibility, mathematical model, infection polymorphism, gene flow, speciation, reproductive isolation barrier, reinforcement, *Drosophila*

Zusammenfassung

Die intrazellulären Parasiten der Bakteriengattung *Wolbachia* sind weit verbreitet im Phylum der Arthropoden. In vielen Wirten lösen sie eine Paarungsinkompatibilität zwischen nicht infizierten Weibchen und infizierten Männchen aus. Die mögliche Rolle dieser zytoplasmatischen Inkompatibilität in Artbildungsprozessen der Wirtsorganismen wird seit langer Zeit diskutiert. In dieser Arbeit analysieren wir häufig angeführte Kritikpunkte einer solchen Rolle mit Hilfe von mathematischen Modellen, in denen Infektionsdynamik von *Wolbachia* und Populationsgenetik der Wirte kombiniert werden.

Die einzelnen Teile befassen sich mit dem Folgenden: (i) Wir untersuchen die Stabilität von Infektionsmustern in Wirts-Metapopulationen, indem wir kritische Migrationsraten herleiten. (ii) Zur Abschätzung des Einflusses der zytoplasmatischen Inkompatibilität auf den Genfluss zwischen Populationen berechnen wir effektive Migrationsraten. (iii) Wir bestimmen die Bedingungen, die die Verstärkung von Reproduktionsbarrieren durch die Evolution von weiblichen Paarungspräferenzen begünstigen. Schließlich (iv) wenden wir unsere Modelle auf einen realen Artbildungsprozess zweier *Drosophila*-Arten in Nordamerika an, diskutieren auftretende Probleme und unterbreiten Vorschläge für weiterführende Forschung.

Zusammenfassend implizieren unsere Ergebnisse, dass *Wolbachien* häufig mit der Entstehung neuer Wirtsarten verknüpft sein köännen, allerdings in den meisten Fällen nur, indem sie als einer von mehreren Faktoren zur reproduktiven Isolation beitragen. Eine Verstärkung sexueller Isolation wird nur unter speziellen Bedingungen bewirkt.

Schlagwörter:
Wolbachia, Zytoplasmatische Inkompatibilität, Mathematische Modellierung, Infektionspolymorphismus, Genfluss, Artbildung, Verstärkung von Reproduktionsbarrieren, *Drosophila*

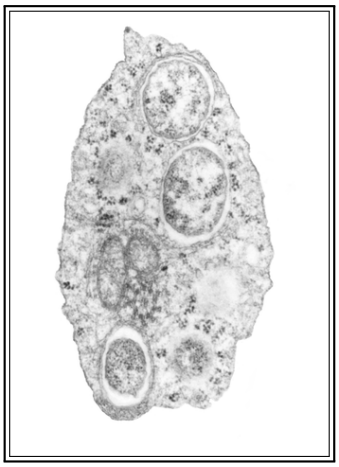
To Saskia and Boyke

Contents

List of Figures

List of Tables

List of Featured Equations



Chapter 1

Introduction

As the sun sets in a beautiful mix of colors at the end of a hot summer day in Berlin, you open the kitchen's bio-waste bin and throw away the left-overs of that delicious dinner you cooked. A few fruit flies (also known as *Drosophila melanogaster*) are stirred up and swirl around. You are probably not aware of it, but chances are that they are infected with a bacterial parasite called *Wolbachia*. These are probably the most abundant endosymbionts in the world (Hilgenböcker et al., 2008), thriving within the cytoplasm of their hosts.

In a lot of insect hosts —including *Drosophila*—, infection with *Wolbachia* can lead to a reproductive incompatibility between infected males and uninfected females, called cytoplasmic incompatibility (Laven, 1951). Now, returning to your kitchen it is quite possible that reproduction between your waste bin's fruit flies and those of a neighbor's down the road is severely impeded due to the incompatibility. This will for example be the case if your population of flies harbors the infection but your neighbor's population does not.

What if uninfected female flies at your neighbor's waste bin start to avoid mating with those infected males that frequently come over from your kitchen? By the end of the summer, reproduction could be even more severely hampered, the cytoplasmic incompatibility now being complemented by female mating preferences. Perhaps, you are just witnessing the rapid emergence of two new fruit fly species ...

Maynard Smith and Szathmáry (1997) list symbiosis as one of three general ways by which genetic complexity may increase during evolution (the others being "duplication and divergence" and "epigenesis"). Most familiar is the endosymbiotic theory for the origin of the eukaryotes (Sagan, 1967), one of the major transitions in evolution: A core feature of eukaryotes are the cell organelles called mitochondria, often referred to as the "powerhouses of the cell". It is now well established that they derive from purple bacteria that were engulfed by heterotrophic prokaryotes and became endosymbionts probably around 2.5 billion years ago. Intriguingly, the ancestors of mitochondria happen to be close relatives of *Wolbachia* (Andersson et al., 1998). Yet, the two have walked entirely different evolutionary paths.

Sharing the cytoplasmic lifestyle of mitochondria, *Wolbachia*'s mode of transmission is not horizontal between unrelated hosts but vertical: the bacteria are inherited and, importantly, solely by females. From the bacteria's point of view, male hosts are dead ends of reproduction and dispersal (Werren, 1997). Manipulations that increase transmission through females at the expense of males will be favored by natural selection. This is the case with the phenomenon called cytoplasmic incompatibility. *Wolbachia* in male hosts modify sperm such that only eggs that harbor clonal relatives of the bacteria (which rescue the modification) can be properly fertilized (Werren, 1997). Uninfected females thus become incompatible with infected males to the relative advantage of infected females, so that *Wolbachia* may spread through a host population (Caspari and Watson, 1959).

Ever since the discovery of cytoplasmic incompatibility, its potential role in speciation has been discussed (Laven, 1959). The basic concept is that, by preventing or severely reducing gene flow between populations, cytoplasmic incompatibility could enhance the probability that populations diverge into separate species (Werren, 1998). Additionally, the evolution of mating preferences could be enforced if uninfected females would thereby decrease their risk of acquiring incompatible mating partners. Such processes run under the label "reinforcement" and are often considered essential steps in speciation (Dobzhansky, 1940).

The "waste bin scenario" described above is of course a fictitious one. We have intentionally connived at several problems. For one, in *D. melanogaster* levels of incompatibility due to *Wolbachia* infections are usually low. More importantly, little is known about infection patterns within host metapopulations in general. Obviously, *Wolbachia* is not to spread nor to go extinct if any effect on speciation is to be expected. Finally, speciation requires more or less complete reproductive isolation, a state that the described female mating preferences certainly would not bring about. Yet, the scenario describes the motivation for this thesis adequately. Our objective is to investigate some aspects (and the stated problems) of such scenarios and the implications for host speciation by means of mathematical models that combine infection dynamics of *Wolbachia* and population genetics of their hosts. In the background, there always lingers this question: To what extent does *Wolbachia*-induced cytoplasmic incompatibility contribute to speciation, or could speciation even be triggered by the incompatibility?

1.1 Overview of this thesis

The present thesis is structured as follows. In Chapter 2, we review the biological context that this thesis is set in. It is divided into two main sections. In the first section, we introduce *Wolbachia* by relating a history of *Wolbachia* research. We then give an account of the striking distribution of these symbionts throughout the arthropod world and of their phylogeny, and describe the bacteria's cell biological and genetical characteristics. Last but not least, we outline the modes of reproductive parasitism that *Wolbachia* engage in, with a special focus on unidirectional cytoplasmic incompatibility (CI)[1].

[1] There is also a *bidirectional* form of cytoplasmic incompatibility which involves two strains of *Wolbachia* that are mutually incompatible.

The second section of Chapter 2 is concerned with speciation by way of reinforcement. We shortly explain the biological species concept that we adhere to in this thesis. It equates speciation to the evolution of reproductive isolation. We then picture the actual process of reinforcement which describes how existing postzygotic isolation enforces the evolution of premating isolation through sexual selection. Finally, we review the possible involvement of *Wolbachia*-induced cytoplasmic incompatibility in speciation, and class the model that we use in this thesis into a general categorization of speciation models.

In Chapter 3, we analyze the conditions that enable the stable coexistence of uninfected populations next to populations infected with *Wolbachia*, a situation we call infection polymorphism. We generalize existing mathematical models of *Wolbachia* infection dynamics to more complex population structures. We use critical migration rates to study mainland-island scenarios, where migration occurs in one direction only, and scenarios with two-way migration. Special cases of the general model are discussed.

Chapter 4 is concerned with the impact that a stable infection polymorphism has on inter-population gene flow. Fitness graphs for the reproductive values of infected and uninfected host organisms are employed to calculate effective migration rates. These measure the reduction in gene flow due to unidirectional CI for both mainland-island scenarios and scenarios with two-way migration.

The process of CI-driven reinforcement is analyzed in Chapter 5. We combine *Wolbachia* dynamics with simple population genetic models to investigate whether unidirectional CI selects for mating preferences that allow females to avoid incompatible matings if an infection polymorphism between populations exists. We analyze the conditions that favor reinforcement of premating isolation and discuss its effect on gene flow. In particular, we determine the influence that costs of female mating preference have on the evolution of premating isolation.

In Chapter 6, we apply our reinforcement model to a real-world speciation example. In North America, the close sibling species *Drosophila recens* (infected with a CI-inducing strain of *Wolbachia*) and *Drosophila subquinaria* (uninfected) show patterns of sexual selection that fit the expected outcome of reinforcement due to unidirectional CI. We inspect how well our model can reproduce the observations.

Finally, we summarize all of our findings in Chapter 7 and discuss future perspectives of our research.

Although the advancement of the leitmotif of CI-assisted speciation can easily be traced throughout this thesis, and despite the fact that each chapter builds on the previous ones, all chapters have been written such that they can be read to a large extent independently of the rest of the thesis. Naturally, this has led to a certain degree of repetition which we hope not to be too irritating.

1.2 Software acknowledgements

In Chapters 3 through 6, we used pencil and paper as well as the huge symbolic computation power of *Mathematica* (Wolfram, 2008) to arrive at analytical results wherever feasible.

Whenever we had to revert to numerical simulations, programs were written in *Python* (`http://www.python.org`), and we made extensive use of the *NumPy*

package (`http://numpy.scipy.org`). *IPython* (`http://ipython.scipy.org`), an enhanced interactive python shell, greatly simplified the process of coding, testing, and plotting (Pérez and Granger, 2007).

Finally, figure plots were edited with *Inkscape* (`http://www.inkscape.org`), and this thesis was written and put together in LaTeX.

1.3 Image acknowledgements

The image that appears above the title of each chapter is a transmission electron micrograph of several *Wolbachia* bacteria within an insect host cell that appeared in a synopsis (PLoS Biology Synopsis, 2004) accompanying the publication of the first complete *Wolbachia* genome by Wu et al. (2004). The image itself is courtesy of Scott O'Neill, published under the Wikimedia Commons free license *cc-by-2.5* at `http://commons.wikimedia.org/wiki/File:Wolbachia.png`.

The phylogeny of *Wolbachia* on page 10 is a slightly edited version of an image by Werren et al. (2008) who in turn used data from Hotopp et al. (2006) and Lo et al. (2007).

The illustrations of female and male *Drosophila melanogaster* that are used in the mating tables (pages 14, 32, 48, 52, and 56) were found at *The Interactive Fly* website (`http://rice.bio.indiana.edu:7082/allied-data/lk/interactive-fly/aimain/1aahome.htm`) but originate from the book *The physical basis of heredity* (Morgan, 1919) by Thomas Hunt Morgan, the first Nobel laureate for genetics.

The phylogeny of the *Drosophila quinaria* species complex on page 123 is adopted from Perlman and Jaenike (2003).

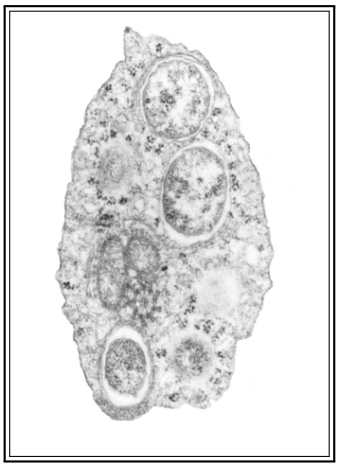
The map on page 124 that shows collection sites of *Drosophila subquinaria* and *D. recens* in North America was taken from Jaenike et al. (2006) (licensed under a Creative Commons Attribution License, see `http://creativecommons.org/licenses/by/2.5/`), converted to greyscale, and supplemented with the presumable geographic ranges of the sibling species.

Permission for reuse in this thesis was acquired for all of the copyrighted images.

1.4 Practical advice

Finally, to give some practical advice on reading this thesis: In the making of the PDF version, hyper references were used to enable simple navigating through the document, i.e., citations, URL's, and numbers of figures, tables, and equations, are all hyperlinks to the content they refer to. In addition, page numbers behind journal articles in the bibliography directly hyperlink to the pages where the article is cited.

Just like the standard *table of contents*, *list of figures*, and *list of tables* simplify finding a particular piece of information, the custom *list of featured equations* on page xi was compiled from the major formulas contained in this thesis and is intended for quickly finding that particular equation.



Chapter 2

Biological background

Bacteria of the genus *Wolbachia* can be considered the most common endosymbionts in the world, infecting many invertebrate species, including insects, spiders, mites, isopods, and nematodes. These α-Proteobacteria are maternally transmitted through the cytoplasm of eggs and have evolved various mechanisms for manipulating reproduction of their hosts. In many insect species, *Wolbachia* induce a sperm-egg incompatibility between the gametes of infected males and uninfected females, so-called cytoplasmic incompatibility.

Speciation according to the biological species concept can be equated to the evolution of barriers to gene flow between groups of organisms that ultimately prevent reproduction between incipient species. It is commonly believed that postzygotic isolation represents an early stage in speciation that can reinforce the evolution of premating isolation. It has been recognized early that cytoplasmic incompatibility can form such a postzygotic isolation barrier between infected and uninfected host populations and therefore might be a factor in parapatric host speciation.

In this introductory chapter, we give an overview of the biology of *Wolbachia* and of speciation via reinforcement.

> The spread of this germline bacterium through the majority of animal species over the last 100 million years represents one of life's great pandemics.
>
> Bordenstein (2008)

2.1 *Wolbachia*

Wolbachia are a remarkable group of bacteria that have attracted growing scientific interest in the last twenty years or so. We will start our introduction into the biology of *Wolbachia* with a brief history of *Wolbachia* research, proceed with a description of the phylogeny and distribution of *Wolbachia*, and finally give an account of the cell biology, genomics and genetics, and host interactions, with a focus on the reproductive parasitism phenotype termed cytoplasmic incompatibility.

First, however, we present a graph that has been created from the scientific citation index at ISI Web of Science, by extracting titles and abstracts from the 100 most cited journal articles on *Wolbachia* (see Fig. 2.1 on the next page). The size of terms corresponds to their frequency in the extracted data. Admittedly not much more than a gimmick, two topics can be seen to dominate research on these bacteria: *cytoplasmic incompatibility* and *infected host species*. This nicely corresponds to the central aspect of the present thesis, namely how *Wolbachia*-induced cytoplasmic incompatibility might be linked to host speciation.

2.1.1 A brief history of *Wolbachia*

Only a few years after da Rocha Lima (1916) had first isolated the etiological agent of typhus fever and named it *Rickettsia prowazeki* after pathologists Ricketts and von Prowazek (both of whom died of typhus), Marshall Hertig and Simeon Burt Wolbach discovered rickettsia-like microorganisms in the reproductive tissue of the mosquito *Culex pipiens* (Hertig and Wolbach, 1924). However, apart from the formal description of the genus as *Wolbachia pipientis* by Hertig (1936), little further investigation was pursued in the following decades.

In the 1950s, Laven (1951, 1956, 1957) and Ghelelovitch (1952) conducted cross-breeding experiments with *Culex* mosquitoes from different locations in Germany, France, and North Africa, and detected complex geographic patterns of reduced egg hatch rates in the crossings. They also established a cytoplasmic inheritance pattern of the causative agent, and Laven (1957) called the phenomenon cytoplasmic incompatibility (CI).

However, it was not before the early 1970s that a connection was first proposed and shortly afterwards proven between this discovery and the rickettsia-like bacteria *Wolbachia* (Yen and Barr, 1971, 1973). Through antibiotic curing of insect hosts, Yen and Barr (1973) were able to restore compatibility of previously incompatible strains, and uncover *Wolbachia* as the etiological agent of CI in *Culex pipiens*.

Until the early 1990s, *Wolbachia* were still considered to be members of a rare bacterial genus, but when polymerase chain reaction (PCR) diagnostics for *Wolbachia* became available (O'Neill et al., 1992; Stouthamer et al., 1993), it became clear that this agent on the contrary was both extremely widespread and also responsible for a range of different reproductive phenotypes in the different hosts it

Figure 2.1: What are *Wolbachia?* This figure gives a graphical answer by showing terms that are commonly associated with *Wolbachia* in the scientific literature. The data used are from the 100 most cited publications that are found by searching for *Wolbachia* in the ISI Web of Knowledge (`http://isiknowledge.com`) ©Thomson Reuters (list of publications retrieved on Jan. 28, 2009), and includes titles, author names, journal names, years of publication, and abstracts. The graph was created on Many Eyes ©IBM (`http://manyeyes.alphaworks.ibm.com/manyeyes/visualizations/terms-linked-with-wolbachia`). The size of a word is proportional to its frequency. Note the strong association of *Wolbachia* with the terms *incompatibility* and *species*, which fairly accurately mirrors the topic of this thesis.

infected. Parthenogenesis-inducing *Wolbachia* were found (Stouthamer et al., 1990, 1993) as well as *Wolbachia* engaging in the feminization of genetic males (Rigaud et al., 1991; Rousset et al., 1992) and *Wolbachia* that selectively kill male offspring (Werren et al., 1994). Werren et al. (1995a) concluded from a screening that at least 16% of neotropical insect species harbored *Wolbachia* infections.

For a long time, the difficulty to culture *Wolbachia* outside of the invertebrate host has been a hindrance to research with these bacteria and has led to a bias toward infections that occur within hosts that are easily reared (e.g., *Drosophila*). However, O'Neill et al. (1997) established in vitro cultivation of *Wolbachia* in host cell lines using standard cell culture techniques. Recently, genome-based metabolic reconstruction has been successfully applied to develop the first cell-free culture medium for the host-dependent bacteria *Tropheryma whippelii* (Renesto et al., 2003). Furthermore, the Metagrowth knowledge resource (`http://igs-server.cnrs-mrs.fr/axenic`) has been installed to collect relevant information to further advance this approach (Ogata and Claverie, 2005). Interestingly, Metagrowth already provides information on the close *Wolbachia*-relative *Rickettsia prowazekii*. If this technique is further established (see also Handorf et al., 2008), cell-free cultivation might well provide the next boost in *Wolbachia* research.

Since Caspari and Watson (1959) and Fine (1978) first modeled *Wolbachia* dynamics in host populations, major advancements in the understanding of *Wolbachia* ecology and population genetics have been made (Hoffmann et al., 1990; Turelli and Hoffmann, 1995; Werren, 1997), including interactions with the spatial structure of host populations (reviewed in Engelstädter and Telschow, 2009) and a potential role of the bacteria in host speciation (see Section 2.2.3).

Rather recently, four complete genomes of different *Wolbachia* strains have become available (Wu et al., 2004; Foster et al., 2005; Klasson et al., 2008, 2009), and a new scheme for typing of *Wolbachia* strains has been developed (Baldo et al., 2006b). We conclude our brief history by noting that in part *Wolbachia* have attracted considerable interest due to their potential application as biocontrol agents (reviewed in Sinkins and Gould, 2006; Floate et al., 2006). *Wolbachia* might be used to directly suppress arthropod populations (Curtis and Adak, 1974; Curtis, 1976; Zabalou et al., 2004), as a vector for the expression of transgenes (Turelli and Hoffmann, 1999; Sinkins and O'Neill, 2000; Rasgon and Scott, 2003), and as a tool to drive desirable genotypes into arthropod populations (Beard et al., 1993; Sinkins and Godfray, 2004). McMeniman et al. (2009) successfully transferred the virulent and life-shortening *Wolbachia* strain *w*MelPop (discovered by Min and Benzer, 1997) into laboratory populations of the mosquito *Aedes aegypti*. This may prove efficient in reducing transmission of dengue viruses which require considerable developmental time in their mosquito vector.

2.1.2 Phylogeny and distribution

In a recent meta-analysis, it was estimated that more than 60% of all insect species are infected with *Wolbachia*, placing it among the most abundant endosymbionts yet discovered and likely infecting more than three million insect species alone (Hilgenböcker et al., 2008). *Wolbachia* have been found in all major insect orders (Diptera, Coleoptera, Hemiptera, Hymenoptera, Orthoptera, and Lepidoptera) as well as in Arachnida and Isopoda (Bourtzis et al., 2003).

Taxon	*Wolbachia*
Domain	Bacteria
Kingdom	Eubacteria
Phylum	Proteobacteria
Class	α-Proteobacteria
Order	Rickettsiales
Family	Rickettsiaceae
Genus	*Wolbachia*
Species	*Wolbachia pipientis*
Supergroup	designated A – H
Strain	e.g. *w*Mel

Table 2.1: Taxonomic classification of *Wolbachia*. Currently, the classification of all strains of *Wolbachia* as belonging to one species (O'Neill et al., 1992) is still valid but far from being settled (Lo et al., 2007). The monophyletic taxon is divided into a number of supergroups. Names of strains typically conform to host organisms, e.g., *Wolbachia* strain *w*Mel is found in *Drosophila melanogaster*.

The taxonomy of *Wolbachia* (summarized in Tab. 2.1) classifies it as one of the closest relatives of the ancestor of mitochondria (Andersson et al., 1998; Wu et al., 2004; Kunisawa, 2005). At present, *Wolbachia* have been divided into eight supergroups (designated A to H) to describe the major phylogenetic subdivisions of this bacterial group (see Fig. 2.2). The designations are primarily based on data from the 16S rRNA, *ftsZ*, and *wsp* (*Wolbachia* surface protein, WSP) genes (Riegler et al., 2005; Werren et al., 1995b; Zhou et al., 1998). Most of the supergroups are found in arthropods (supergroups A, B, E, F, G, and H), supergroups C and D are associated with filarial nematodes, and the majority of insect *Wolbachia* strains belong to supergroups A and B (Baldo et al., 2006b). The phylogenetic relationship between these supergroups is currently not well resolved (Lo et al., 2007).

A major advancement for the study of *Wolbachia* diversity was achieved when Baldo et al. (2006b) developed a multilocus sequence typing (MLST) scheme as a universal genotyping tool for *Wolbachia*. Their scheme is based on internal fragments of five ubiquitous genes (*gatB*, *coxA*, *hcpA*, *fbpA*, and *ftsZ*) that is supplemented by an additional typing system using the hypervariable regions of the *wsp* gene. Baldo et al. (2006b) found a total of 35 unique allelic profiles, and this number has grown to more than 180 up until the writing of this thesis. All profiles and strains can be accessed at the public *Wolbachia* MLST database (`http://pubmlst.org/wolbachia`). An earlier effort to document the great diversity of this bacterial genus has been undertaken by the *Wolbachia* host database (`http://www.wolbachia.sols.uq.edu.au`), where 450+ strains are listed (Riegler and O'Neill, 2006), cross-linked with GenBank (`http://www.ncbi.nlm.nih.gov/GenBank`) of the National Center for Biotechnology Information (NCBI). Typing of these strains however varies (primarily, data from 16S rRNA, *ftsZ*, or *wsp* were used) and is thus not consistent.

The lack of congruence of bacterial and host phylogenies clearly shows that *Wolbachia* rarely cospeciate with their hosts, and that intertaxon, or horizontal transmission must have occurred frequently in the evolutionary past (Werren, 1997; Zhou et al., 1998; Vavre et al., 1999; Baldo et al., 2006a). There also is evidence of extensive recombination between *Wolbachia* strains and even supergroups (Baldo et al., 2006a). The apparent short evolutionary persistence times of *Wolbachia* in many host systems may be due to the spread of host modifier alleles (Koehncke et al.,

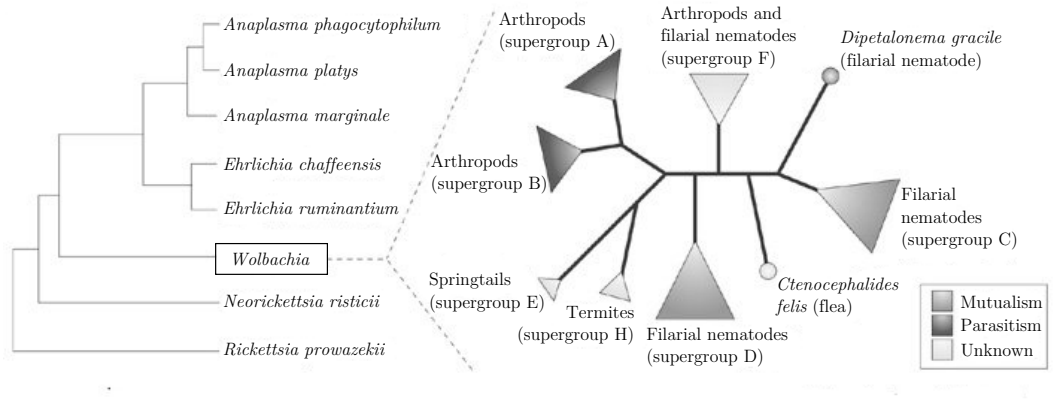


Figure 2.2: Phylogeny of *Wolbachia*. The left part shows *Wolbachia*'s phylogenetic relationship relative to the closely related Rickettsiales order in the Anaplasmataceae family. In the right part, an unrooted phylogenetic tree of *Wolbachia* supergroups A–H is shown (note that the G supergroup has been removed because its status is currently unclear). Supergroups seem to participate either in mutualism or reproductive parasitism. Rooting of the tree is problematic, so that it is currently unresolved which relationship is ancestral. Triangle size represents described diversity within each linage, circles mark lineages based on a single *Wolbachia* strain. Image and legend adopted from Hotopp et al. (2006); Lo et al. (2007); Werren et al. (2008).

2009) that eventually result in a host population losing the infection. The actual mode of horizontal transmission, however, remains yet to be determined (but see the next Section 2.1.3 for some recent findings).

2.1.3 Cell biology, genetics, and host interactions

Wolbachia are Gram-negative α-Proteobacteria that live as obligate intracellular endosymbionts. Being coccoid or bacilloform in morphology and measuring about 0.8–1.5 μm in length (Hertig, 1936), they are located inside a vacuole, i.e., they are surrounded by three membranes: an inner plasma membrane, an outer cell wall, and the vacuole membrane that is of host origin (Louis and Nigro, 1989, also cp. the electron micrograph image above each chapter title). The extent to which this outermost membrane is modified by the bacteria is yet to be determined (Riegler and O'Neill, 2006), although genomic analysis indicates that *Wolbachia* have lost essential parts of the usual machinery to build cell walls (Wu et al., 2004).

The bacteria are predominantly localized in the ovaries and testes of their hosts but can also be detected in somatic tissues such as muscle or nervous cells (Dobson et al., 1999). Because of their intracellular habitat, *Wolbachia* can not be easily preserved. Cultivation in cell-free medium is not available yet although Rasgon et al. (2006) have demonstrated the bacteria's ability to survive outside of host cells. However, as mentioned above, maintanance of *Wolbachia* in cell lines (that can be stored at –80 °C) is well established (O'Neill et al., 1997; Dobson et al., 2002b).

Over ecological timescales, *Wolbachia* infections are maintained by strict maternal inheritance (Stouthamer et al., 1999) through the female germline, with transmis-

sion efficiency commonly close to 100% (Hoffmann et al., 1998, 1990; Turelli and Hoffmann, 1995). However, as reported above, comparative phylogenetics indicate that over evolutionary timescales, horizontal transfer of *Wolbachia* must be common, resulting in closely related infections in very distantly related hosts (Zhou et al., 1998; Stouthamer et al., 1999). The occurrence of horizontal transfer is further underpinned by the fact that *Wolbachia* strains can be artificially transferred into novel hosts by embryonic microinjection (Boyle et al., 1993; Braig et al., 1998). This enables studies of host genetic background effects on *Wolbachia* phenotypes (e.g., Clancy and Hoffmann, 1997; Poinsot et al., 1998). Also by way of microinjection, Rousset et al. (1999) were able to establish a stable triple *Wolbachia* infection in *Drosophila* flies. Frydman et al. (2006) showed that *Wolbachia* bacteria are able to migrate through somatic tissues into the germline via the somatic stem cell niche.

Wolbachia are often found in close association with microtubules or the cell cortex. In late oogenesis, the bacteria localize at the posterior end of the oocyte and thus promote their germline-based transmission mode (e.g., Breeuwer and Werren, 1990; Stouthamer et al., 1993; Serbus and Sullivan, 2007). It has been demonstrated that *Wolbachia* utilize the host microtubule network and associated proteins for their subcellular localization in the *Drosophila* oocyte (Ferree et al., 2005; Serbus and Sullivan, 2007).

During spermatogenesis, *Wolbachia* are predominantly localized at the proximal end of the immature cyst, opposite the spermatid nuclei, and are only excluded from mature sperm during the final stages of sperm maturation (Clark et al., 2002; Hoffmann et al., 1998; Turelli and Hoffmann, 1995). Males therefore are evolutionary dead-ends of inheritance for *Wolbachia*, and this sets the stage for the various forms of reproductive parasitism that *Wolbachia* have become notorious for: Feminization of genetic males, induction of Parthenogenesis, male-killing, and cytoplasmic incompatibility. All of these *Wolbachia*-host interactions increase —under appropriate circumstances— the relative fitness of infected females and thus foster the propagation of the bacteria; they are discussed in more detail below.

In nematodes (Taylor and Hoerauf, 1999), and potentially in the parasitic wasp *Asobara tabida* (Dedeine et al., 2001, 2005) and the mosquito *Aedes albopictus* (Dobson et al., 2002a), *Wolbachia* have evolved a mutualistic relationship with their hosts. The exact benefits of the bacteria to their hosts have not been determined, but antibiotic curing results in premature death of worm larvae or failure of wasp ovaries to develop properly. Recently, Weeks et al. (2007) reported the rapid evolution of mutualistic behavior of *Wolbachia* in *D. simulans* (however, without change to their parasitic mode of cytoplasmic incompatibility). In another study, it was suggested that *Wolbachia*-infection might provide resistance against RNA viruses in *Drosophila* (Teixeira et al., 2008; Osborne et al., 2009). Despite considerable advancements in understanding the cytology of *Wolbachia*-host interactions (for a review of the current knowledge, see Serbus et al., 2008), the underlying molecular basis remains to a large extent unknown.

To our best knowledge, fully annotated genomes of four *Wolbachia* strains are available. Three of them are CI-inducing strains: *w*Mel, a supergroup A strain from *Drosophila melanogaster* (Wu et al., 2004), *w*Pip, a supergroup B strain infecting the mosquito *Culex quinquefasciatus* (Klasson et al., 2008), and *w*Ri, a supergroup B strain naturally infecting *D. simulans* (Klasson et al., 2009). The fourth strain,

*w*Bm, belongs to supergroup D and is an obligate mutualist of its nematode host *Brugia malayi* (Foster et al., 2005). Several other *Wolbachia* genome projects are underway (see Table 1 in Werren et al., 2008). *Wolbachia* genomes are small (1.08–1.7 Mb) but lack the typical minimal content common in other obligate endosymbionts (Werren et al., 2008). Mobile and repetitive elements —predominantly ankyrin (ANK) domains— account for considerable parts of *Wolbachia* genomes (Wu et al., 2004). ANK repeats are known to play a role in cell cycle regulation (Elfring et al., 1997). Some evidence suggest that these repeats, together with virus-like elements (such as the lambda bacteriophage WO) which are also present in large numbers at least in the *w*Mel genome, are involved in *Wolbachia*–host interactions, including reproductive manipulations (Sinkins et al., 2005; Walker et al., 2007). Recently, Hotopp et al. (2007) found evidence for extensive lateral gene transfer: In *Drosophila ananassae* Hawaii, nearly a complete *Wolbachia* genome was transferred into the fly nuclear genome.

2.1.4 Reproductive parasitism

With maternal transmission being imperfect, or the infection having negative effects on female fecundity, *Wolbachia* would inevitably become lost from current populations if not for their manipulation of host reproduction. All of these manipulations produced by *Wolbachia* bacteria lead to a relative increase in the number of surviving daughters produced by infected individuals and thus enable the persistence or spread of the infection. Four modes of manipulations have been found, three of which represent different forms of sex ratio distortion. *Feminization*, in which infected genetic males develop and reproduce as females, and *induction of parthenogenesis*, in which infected virgin females produce daughters only, have straightforward beneficial effects from the bacteria's point of view. If *male-killing*, in which infected sons die at an early embryonic stage, is to be advantageous for *Wolbachia*, surviving infected daughters must benefit from their brothers' death through fitness compensation. The perhaps most surprising and intriguing manipulation (and the only one without sex ratio distortion) is called *cytoplasmic incompatibility* (CI), in which sperm of infected males is reproductively incompatible with uninfected eggs. Here, the bacteria's advantage is less transparent and more indirect because it is the progeny of uninfected females that is (negatively) affected.

CI is probably the most common phenotype (see the section on incidences of the different phenotypes on page 16); and as it is central to this thesis, it will be presented in more detail than the others. Note that male-killing and cytoplasmic incompatibility both can only be understood in terms of kin selection as it is not the "trapped" bacteria themselves that benefit from the manipulations but their clonal relatives.

Only one other bacterium, *Cardinium* of the phylum Bacteroidetes, has been found to have similarly diverse effects on host reproduction (Zchori-Fein et al., 2004). *Cardinium* is also the only other known bacterial inducer of cytoplasmic incompatibility so far.

Feminization. Sex ratio distortion in favor of females is common in terrestrial isopods (woodlice) which exhibit female heterogametic sex-determination (males

ZZ and females ZW); and feminization of genetic males through cytoplasmic microorganisms has been long known to occur in *Armadillidium vulgare* (Martin et al., 1973). Rigaud et al. (1991) were the first to identify the causative agent as *Wolbachia*. The bacteria proliferate within the androgenic gland, leading to a hypertrophic and non-functional gland. Consequently, genetic males develop as females (Juchault et al., 1994; Vandekerckhove et al., 2003). Feminizing *Wolbachia* have also been found in insects where the bacteria supposedly interfere with the sex-determination pathway of their hosts (Hiroki et al., 2002). Their presence throughout host development is necessary to complete feminization (Werren et al., 2008). It has been suggested that the nuclear-cytoplasmic conflict over sex-determination caused by this *Wolbachia* phenotype can provoke rapid changes of sex-determination systems (Rigaud and Juchault, 1993; Rigaud et al., 1997).

Parthenogenesis induction. In arrhenotokous species in which haploid males develop from unfertilized and diploid females from fertilized eggs, such as mites or hymenopterans, *Wolbachia* can induce parthenogenesis (Huigens and Stouthamer, 2003). In these haplodiploid sex-determination systems, *Wolbachia* disrupt the first mitotic division in unfertilized eggs so that gametic duplication yields diploid nuclei that are homozygotic at all loci (Stouthamer et al., 1990; Stouthamer and Werren, 1993; Stouthamer and Kazmer, 1994). Usually, antibiotic curing restores production of fertile males, but in the parasitic wasp *Encarsia formosa*, cured males fail to inseminate females to the effect that *Wolbachia* might have irreversibly converted an entire species from sexual to asexual reproduction (Zchori-Fein et al., 1992).

Male-killing. Male-killing (MK) is in some sense unusual among the reproductive manipulations exhibited by *Wolbachia* in being a commonly found phenotype within Eubacteria in general (reviewed in Hurst and Jiggins, 2000). *Wolbachia*-induced MK has been described in four different arthropod orders (Werren et al., 2008): Coleoptera, Diptera, Lepidoptera, and Pseudoscorpions. In nearly all instances, *early* male-killing, in which infected male embryos are killed at an early stage, is produced. In this case, reallocation of freed resources to infected daughters benefit clonal relatives of the male-killer and result in spread of the infection[1] (Hurst, 1991).

The prevalences of MK-*Wolbachia* infections range from just above 0% in many *Drosophila* species to almost 100% in some butterflies (Hurst and Jiggins, 2000). Consequently, male-killing sometimes causes extreme sex-ratio biases in favor of females in natural populations with the possible extermination of the populations due to lack of males (see, e.g., Jiggins et al., 2000; Dyson and Hurst, 2004). Hornett et al. (2006) recorded the rapid spread of a host mutation that suppresses MK in the butterfly *Hypolimnas bolina*. Interestingly, the suppression unveiled a new phenotype: males with the suppressor allele survive to adulthood but when mated to uninfected females, they induce CI (Hornett et al., 2008). Along the same lines, Jaenike (2007) described the rapid shift of a CI phenotype to an MK phenotype upon introgression of *Wolbachia* into a novel host (see also Chapter 6). Furthermore, the feminization phenotype that *Wolbachia* was long thought to induce in

[1] *Late* male-killing, in which fourth instar larvae are killed, has a different evolutionary logic: It is associated with microsporidians which gain horizontal transmission after having killed the male progeny.

the lepidopteran host *Ostinia scapulalis* has been found to be male-killing instead (Kageyama and Traut, 2004). In the absence of the bacteria, females die during larval development. If *Wolbachia* is present, genetic males develop as females but die in larval stages.

The tight connections between the phenotypes, apparent from these examples, further raise the questions as to how similar the underlying molecular mechanisms possibly are, how important the host genetic background is for the expression of a phenotype, and how widespread multi-potent *Wolbachia* might be that possess the machinery for inducing multiple phenotypes (Werren et al., 2008).

Cytoplasmic incompatibility. Cytoplasmic incompatibility (CI) is generally believed to be the most common form of reproductive parasitism enforced by *Wolbachia* —common here meaning both *frequent* and *widely distributed* (also see the section on phenotype incidences on page 16). The first notion of an incompatibility between strains of the mosquito *Culex pipiens* was given by Marshall (1938). As mentioned above, it was Laven (1951) and Ghelelovitch (1952) who established cytoplasmic inheritance of the incompatibility factor, and Yen and Barr (1973) provided evidence that *Wolbachia* is the causative agent of cytoplasmic incompatibility.

In general, unidirectional CI is a reproductive incompatibility between sperm of *Wolbachia*-infected males and eggs from uninfected females (see Tab. 2.2). Offspring in these matings suffers high mortalities usually in early zygotic stages, whereas the reciprocal mating between an uninfected male and an infected female produces normal numbers of viable and fertile offspring. Hence the incompatibility is termed unidirectional. Bidirectional CI can occur in matings between host populations that are infected with different *Wolbachia* strains. For example, in Laven's original experiments (Laven, 1957, Fig. 1), unidirectional CI is observed when *Culex* females

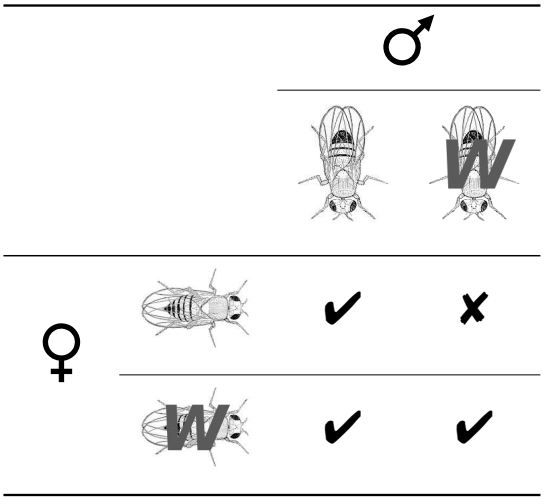


Table 2.2: Cytoplasmic incompatibility (CI). CI-inducing *Wolbachia (W)* evoke a mating incompatibility between uninfected females (♀) and infected males (♂). In these matings, less or no offspring are produced (✘). All other pairings are compatible and yield normal offspring numbers (✔). The table shows unidirectional CI. If mating partners are infected with different strains of *Wolbachia*, bidirectional CI may occur.

from Oggelshausen (Southern Germany) fail to produce offspring when mated to males from Hamburg (Northern Germany) whereas Oggelshausen males can produce offspring with Hamburg females. An example of bidirectional CI is evident in matings between the *Culex* populations from Oggelshausen and Scauri (close to Rome, Italy).

Werren (1997) described CI as a modification-rescue system: *Wolbachia* within an infected male modify sperm. Following fertilization of a *Wolbachia*-free egg, the modification leads to developmental failure. If however, the same (or a similar enough) strain of *Wolbachia* is present in the egg, the modification is rescued by the bacteria, and development proceeds normally. Importantly, the rescue factor alone (i.e., if an infected egg is fertilized by unmodified sperm) does not have negative effects and is compatible with normal development.

The fact that CI is widespread in arthropods and that it causes the same phenotype both at the chromosomal and organismal levels in a wide variety of arthropods suggests that the bacteria interfere with fundamental, but conserved, molecular and developmental processes (Stouthamer et al., 1999). Cytological studies of the CI mechanism have been conducted predominantly in *Drosophila* and in the parasitic wasp *Nasonia*. CI produces dramatic defects in the first mitotic division of the zygote but developmental anomalies also occur at later stages (Callaini and Riparbelli, 1996). One of the earliest manifestations of CI is the delay of both the breakdown of the paternal nuclear envelope and of Cdk1 (cyclin dependent kinase 1) activation relative to the female pronucleus in prophase, prior to DNA condensation (Tram and Sullivan, 2002). This might indicate that *Wolbachia* directly targets cell cycle regulation/checkpoints. The finding, however, that in *Drosophila* androgenetic development (i.e., from the paternal pronucleus only) is inhibited by the *Wolbachia*-induced modification argues against pronuclear asynchrony as the primary cause of CI lethality (Ferree and Sullivan, 2006). In metaphase, male chromatids are still in a semi-condensed state when female chromatids are already properly condensed (Reed and Werren, 1995). Consequently, a hallmark of incompatible crosses is the presence of chromatin bridges between the female and male nuclei in anaphase (Serbus et al., 2008). An all-important feature of CI is the rescue of the sperm modification in an infected egg. Accordingly, the described timing and condensation defects are suppressed by the rescue function.

Several models have been proposed to explain the observed modification-rescue, or *mod-resc*, system of CI (for a comparison of these models, see Poinsot et al., 2003; Bossan, 2009). The lock-key hypothesis proposes that *Wolbachia* in the male host produce a "lock" (the *mod* function) that binds to the paternal nucleus and causes mitotic defects unless bacteria in the egg cytoplasm provide "keys" for unlocking and restoring development (the *resc* function). The titration-restitution hypothesis (Kose and Karr, 1995; Reed and Werren, 1995) takes the reverse logic and states that *Wolbachia* in the male might remove an essential component from the male pronucleus (*mod*) and maternal *Wolbachia* then would restore this factor (*resc*). In both the lock-key and the titration-restitution model, *mod* and *resc* are independent functions that can therefore also evolve independently. By contrast, the third proposed model –the mistiming hypothesis— allows for *mod* and *resc* to be essentially the same factor or function. It suggests that *Wolbachia* produce a factor binding to the male pronucleus and slowing it down (*mod*) to the effect of

asynchrony between male and female pronuclei (Callaini et al., 1997; Tram and Sullivan, 2002). In an infected egg, the maternal chromosomes would be similarly modified (*resc*), and thus synchrony would be restored. The hypotheses have to be tested against the ever accumulating amount of experimental findings, and the final word on what hypothesis explains these data best has not been spoken (Poinsot et al., 2003).

While in most host species the strength of CI correlates with the bacterial load of sperm cysts in the testis (Clark et al., 2003; Veneti et al., 2003), in some hosts this correlation is absent (Duron et al., 2007). Furthermore, in *D. simulans* sperm modification was shown to be effective even in uninfected cysts which suggests a diffusible CI-inducing factor that can spread from infected to uninfected cells throughout the testis (Riparbelli et al., 2007). Yet, Presgraves (2000) elegantly showed through the use of gynogenetic females that do not need an actual male pronucleus but still rely on extranuclear paternal factors that the modification induced by *Wolbachia* affects the male pronucleus itself and no extranuclear component of the sperm. The level of incompatibility has been shown to decline with male age (Duron et al., 2007) and with slower male development (Yamada et al., 2007).

Genetically, the most promising candidates for the study of *Wolbachia*'s phenotypes are ankyrin repeats, WO prophage genes, and *wsp* which contains four hypervariable regions (Werren et al., 2008). However, to date only weak correlations between either of these groups of genes and CI patterns have been found (Sinkins et al., 2005; Duron et al., 2006b; Serbus et al., 2008).

In their review of cytoplasmic incompatibility, Bourtzis et al. (2003) list a large number of studies in which *Wolbachia*-induced CI was reported, spanning all major insect orders (Diptera, Coleoptera, Hemiptera, Hymenoptera, Orthoptera, and Lepidoptera), as well as two families of the class Arachnida and one family of Crustacea. Based on the distribution of CI within the general phylogeny of *Wolbachia*, Bourtzis et al. (2003) state that it is evident that CI is the most frequent and widely distributed phenotype of *Wolbachia*, and that it is "parsimonious to assume that [CI] was an ancestral *Wolbachia* property".

Incidences. While in virtually every journal article on *Wolbachia*, CI is stated as the most frequent and widely distributed phenotype, just how common the different *Wolbachia*-induced phenotypes really are, can still be considered a matter of debate. This is evident from the phenotype frequencies among the strains recorded in the *Wolbachia* MLST database (`http://pubmlst.org/wolbachia`) and in the *Wolbachia* host database (`http://www.wolbachia.sols.uq.edu.au`) that are presented in Tab. 2.3. It is obvious, however, that the knowledge is still very limited. It should be kept in mind, that as maternal transmission of *Wolbachia* in general is imperfect, some mechanism must be at work in every infected host species that maintains the infection, be it a mutualistic or a parasitic one.

As stated above, Bourtzis et al. (2003) list a large number of studies on a diverse range of hosts to support the claim of CI being the most common manipulation induced by *Wolbachia*. However, the count for the other phenotypes is far from conclusive. Hurst et al. (2003) find that certain insect families are particularly prone to male-killer infections due to their ecology and behavior. They conjecture that there might be more than 500 *Wolbachia*-infected Coccinellidae species alone.

Phenotype	MLST db	Host db
Feminization	3	10
Parthenogenesis induction	3	25
Male-killing	5	1
Cytoplasmic incompatibility	28	37
Unknown	243	400+

Table 2.3: Frequencies of reproductive phenotypes. As of this writing, there are 282 records in the *Wolbachia* MLST database and more than 450 in the *Wolbachia* host database (for web addresses, see main text). For most strains, the associated phenotype is unknown. However, among the strains with known phenotype, CI is the most common.

Furthermore, Huigens and Stouthamer (2003) summarize that 46 hymenopteran species have been found to be infected with parthenogenesis-inducing *Wolbachia* and expect that the actual incidence will be much higher than this number, partly because of low infection frequencies within host populations that easily go undetected. Similar conclusions were drawn by Bouchon et al. (1998) who in a screening of isopod crustaceans found twenty-two out of 85 species infected and likely affected by intersexual development or feminization. Generally, it is hard to tell how much the phenotype incidences reported in screenings are biased due to previous knowledge of *Wolbachia*-host systems. In summary, the numbers presented in Tab. 2.3 might change considerably with the advent of further research results.

> Species are groups of interbreeding natural populations that are reproductively isolated from other such groups.
>
> Mayr (1969, p. 314)

2.2 Speciation

The origin of species is one of the central themes in evolutionary biology, the unifying theory of life sciences founded by Charles Darwin. "The point has often been made that, despite its title, [Charles Darwin's] *The Origin of Species* (1859) had much more to say about change *within* species than about the origin of new species", as Coyne and Orr (2004, p. 1, emphasis in the original) have put it in the introduction to their instant classic book on speciation. This was partly due to the difficulty in defining what a species is; but apparently, Darwin was not even sure whether species were real or rather theoretical constructs of the human mind (Mayr, 1982). The great majority of biologists, however, does accept the reality of species, being apparent from the discontinuities of the living world (Dobzhansky, 1937). The Modern Synthesis provided a widely accepted species definition (although not undisputed to date; for a review of the competing concepts, see the appendix in Coyne and Orr, 2004), known as the *biological species concept* (Mayr, 1942). For the purposes of the present thesis, it also enables us to study the process of speciation.

2.2.1 Biological species concept and reproductive isolation

The Ernst Mayr quotation in the epigraph above dates from 1969. The original definition of the biological species concept (BSC) read "species are groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups" (Mayr, 1942, p. 120). Because of the difficulties of determining the species status of allopatric taxa, Mayr dropped the word "potentially" from his definition. Most evolutionary biologists adhere to the BSC, and we will do so in the present thesis likewise. We will albeit adopt a slightly different version put forth by Coyne and Orr (2004, p. 30) which can be formulated in short as follows:

> Species are groups of interbreeding natural populations that are *substantially* reproductively isolated from other such groups.

In the light of evidence that hybridization and introgression occur with considerable frequencies (e.g., Powell, 1983) so that species are not necessarily completely isolated, this definition appears more appropriate. In any case, the exact definition is less important than to recognize that the process of speciation involves acquiring reproductive barriers between emerging species.

Reproductive isolation can take many forms. An obvious subdivision of reproductive isolation barriers can be made with respect to the life cycle of the interbreeding organisms (see the classification in Tab. 2.4): *premating isolation barriers* impede gene flow before transfer of sperm to members of other populations, *postmating/prezygotic isolation barriers* act after sperm transfer but before fertilization, and *postzygotic isolation barriers* act after fertilization.

I. **Premating isolation barriers:** Impediment of gene flow before transfer of sperm or pollen to members of other species.
   A. **Behavioral isolation:** Also called "sexual isolation"; prevention of courtship initiation or copulation due to lack of cross-attraction.
   B. **Ecological isolation:** Isolation as byproducts of adaptation to local environments; can be further subdivided into **(i) habitat isolation**, **(ii) temporal isolation**, and **(iii) pollinator isolation**.
   C. **Mechanical isolation:** Incompatibility of reproductive structures.
   D. **Mating system "isolation":** Evolution of autogamy or apomixis; not considered an isolation barrier in the same sense as the others in this list.

II. **Postmating, prezygotic isolation barriers:** Isolation acting after sperm or pollen transfer but before fertilization.
   A. **Copulatory behavioral isolation:** Behavior during copulation is insufficient to allow normal fertilization.
   B. **Gametic isolation:** Transferred gametes cannot effect fertilization; further subdivision into **(i) noncompetitive** (intrinsic problems) and **(ii) competitive** (problems only in the presence of conspecific gametes).

III. **Postzygotic isolation barriers (hybrid sterility and inviability):** Isolation acting after fertilization.
   A. **Extrinsic:** Dependence on the environment; further subdivision into **(i) ecological inviability** (hybrids develop normally but suffer lower viability because they cannot find an appropriate niche) and **(ii) behavioral sterility** (hybrids have normal gametogenesis but are less fertile because they are unattractive to parent species members).
   B. **Intrinsic:** Developmental problems relatively independent of the environment; subdivision into **(i) hybrid inviability** (fully or partially lethal developmental failures) and **(ii) hybrid sterility** (physiological or behavioral).

Table 2.4: Classification of reproductive isolation barriers. The classification presented here is a slightly edited version of the list by Coyne and Orr (2004, pp. 28–29) who updated the classical one by Dobzhansky (1937, pp. 231–232) in the light of recent work.

The relative importance of an isolation barrier follows directly if one considers that a later-acting barrier will only reduce gene flow that has escaped earlier-acting barriers (Coyne and Orr, 2004). For example, if we assume n isolating barriers acting sequentially over the life cycle with respective absolute strengths $0 \leq I_i \leq 1$, then the total gene flow between a pair of species, G, is

$$G = (1 - I_1)\,(1 - I_2) \cdots (1 - I_n) = \prod_{i=1}^{n} (1 - I_i)\,. \tag{2.1}$$

Thus, an isolation barrier which is strong in absolute terms need not be important in minimizing hybridization if it acts very late in the life cycle.

One implication of this argument is that the reproductive barriers currently most important in restricting gene flow need not be those historically most important during speciation (Coyne and Orr, 2004). For instance, the present inviability of hybrids between the sympatric *Drosophila* species *D. simulans* and *D. melanogaster* is caused by around 190 hybrid-lethal genes, each of which causes nearly complete inviability on its own when in a hybrid background (Presgraves, 2003). But surely, most of these incompatibilities have accumulated after speciation was already completed, so it is difficult to say which genes were involved in the actual process of speciation. Thus, the order in which reproductive barriers occur during speciation processes is important if one is to understand what evolutionary forces drive speciation.

From the relative contributions of sequentially acting isolation barriers, it is furthermore straightforward to deduce that prezygotic isolation barriers are likely to be more effective than postzygotic barriers such as hybrid incompatibility in maintaining species differences despite gene flow (Kirkpatrick and Ravigné, 2002; Coyne and Orr, 2004).

2.2.2 Sexual selection and reinforcement

Given enough time, speciation is an inevitable consequence of populations evolving in allopatry, simply because no forces act on maintaining reproductive compatibility between geographically isolated populations (Mayr, 1942; Turelli et al., 2001). Thus, they will eventually become reproductively isolated. By contrast, in parapatric and sympatric speciation, gene flow between populations is on-going and constitutes exactly such a force that acts on maintenance of compatibility. Consequently, one has to explain how divergence evolves despite gene flow.

The theory of sexual selection advanced by Darwin (1874) was intended to explain the common occurrence of striking secondary sexual characters which cannot easily be attributed to natural selection on viability or fecundity. Fisher (1930) elaborated on the subject, most notably developing a verbal model of runaway selection in which the fitness of a male trait with respect to mating success due to female mating preferences can override its value for survival. He cites male bird plumage as an example:

> "The importance of this situation lies in the fact that the further development of the plumage character will still proceed, by reason of the advantage gained in sexual selection, even after it has passed the point in development at which its advantage in Natural Selection has ceased. [...] Moreover, as long as there is a net advantage in favour of further plumage development, there will also be a net advantage in favour of giving to it a more decided preference. The two characteristics affected by such a process, namely plumage development in the male, and sexual preference for such developments in the female, must thus advance together, and so long as the process is unchecked by severe counterselection, will advance with ever-increasing speed. [...] There is thus in any bionomic situation, in which sexual selection is capable of con-

> ferring a great reproductive advantage, the potentiality of a runaway process, which, however small the beginnings from which it arose, must, unless checked, produce great effects, and in the later stages with great rapidity." (Fisher, 1930, pp. 136–137)

At about the same time, Dobzhansky (1940) argued that changes in species recognition systems could evolve owing to direct selection to avoid deleterious hybridization; he termed this evolution of prezygotic isolation due to existing postzygotic reproductive barriers *reinforcement* and considered it to be the finalizing step in speciation. However, during the Modern Synthesis, these processes of mate choice and of species recognition were seen as fundamentally different (Ritchie, 2007). It was only in the 1980s that the continuity of sexual preferences from *within populations* to *between populations* became more emphasized, and a connection to speciation was drawn (Lande, 1981; West-Eberhard, 1979). Since the 1990s, a significant upturn in the frequency of papers published on sexual selection and speciation can be observed (Ritchie, 2007).

Lande (1981) in quantitative polygenic models and Kirkpatrick (1982) in a conventional population genetic model formalized the verbal reasoning of Fisher (1930) and showed that as long as there is variation for a male trait and a female mating preference for that trait, genetic covariance between the two results, producing a neutrally stable line of equilibria along which the system can move by genetic drift. In Lande's (1981) model, this line of equilibria can become unstable if the linkage disequilibrium between trait and preference is particularly strong, in which case the system evolves rapidly away from the line of equilibria in an unpredictable direction, thus enabling fast divergence between populations.

However, as Bulmer (1989) pointed out, both models are structurally unstable. If either mutation or weak direct selection on female preferences are introduced, the line of equilibria collapses to a single equilibrium point. Genetic drift thus is not expected to be a relevant factor in population divergence by reinforcement. Along similar lines, Iwasa and Pomiankowski (1995) showed that, depending on a mutation bias acting on the male trait and costs of female mating preferences, Fisherian runaways can yield cyclic evolution of male trait and female mating preferences (small bias and costs) or attainment of a stable equilibrium (large bias and/or costs). They suggest that if evolution is cyclic then allopatric populations will quickly fall out of phase and evolve distinct sexual phenotypes.

The most straightforward explanation of female preferences for male traits, however, is that they are either directly selected for (for instance, due to selection against hybrids) (Kirkpatrick and Ryan, 1991) or that they are pleiotropic side effects of alleles selected for other reasons (such as selection for finding prey or mates) (Price, 1998; Turelli et al., 2001). In fact, most recent models and empirical studies of reinforcement support the assumption that some form of direct selection on female mating preferences must be involved (e.g., Kirkpatrick and Servedio, 1999; Noor, 1999; Hall et al., 2000; Servedio, 2000).

Servedio and Noor (2003) in the most recent review on the topic conclude that it is no longer a question of whether reinforcement can or does happen but rather how frequent and important a factor it is relative to other means of speciation. Interestingly, they suggest that researchers have not looked for reinforcement in the most promising systems yet. For instance, both theoretical and empirical studies

have concentrated almost exclusively on intrinsic hybrid incompatibilities rather than extrinsic incompatibilities in which lower fitness is associated with a specific set of ecological conditions (Coyne and Orr, 1998). In this light, it is interesting to note that the classification of a unidirectional CI barrier between populations with different infection statuses is somewhat ambiguous, which we will discuss in the next section.

2.2.3 Cytoplasmic incompatibility and speciation

Current estimates of the total number of species on Earth range from three to thirty million or more, with arthropod diversity contributing a large part (Erwin, 1982; May, 1988; Ødegaard, 2000). This species richness among arthropods is astounding. Several broader patterns of postzygotic isolation —the commonness of hybrid sterility, sex-limited inviability, and "speciation clocks"— rather seem to deny *Wolbachia* a ubiquitous role in arthropod speciation (Coyne and Orr, 2004). However, given the fact that *Wolbachia* is so widespread among this animal phylum, one can but ask for a potential connection, i.e., could *Wolbachia* be a frequent factor in arthropod speciation and thus at least add to this extraordinary species diversity?

As mentioned in the previous section, the classification of the reproductive isolation barrier imposed by unidirectional CI according to Tab. 2.4 is ambiguous. It has been stated that CI poses an intrinsic postzygotic isolation barrier (Coyne and Orr, 2004). Here, we argue that there are extrinsic components to the isolation barrier as well because CI is frequency-dependent to the effect that viability and fertility of hybrids strongly depend on the population's infection status. As Tab. 2.5 shows, the effects on reproductive isolation certainly become more complex than a simple intrinsic factor if one assumes incomplete cytoplasmic incompatibility. Furthermore, there are sex-specific effects which originate from the asymmetrical nature of CI (cp. Tab. 2.2).

Bordenstein (2003) lists four modes of CI-assisted speciation. First, bidirectional CI alone can limit gene flow and maintain genetic divergence between populations (Telschow et al., 2002a), in which case new species could arise in the absence of nuclear divergence. Phylogenetic evidence suggests that such a scenario requires two allopatric populations to independently acquire infections with two incompatible *Wolbachia* strains which on secondary contact would prevent hybridization. It can however be argued that this probably does not occur very frequently in nature; i.e., it seems unlikely that *Wolbachia* plays a major role in host speciation through this mode.

Second, CI (in both its bi- and unidirectional form) can act together with other isolation barriers to further reduce gene flow between incipient species. In the case of unidirectional CI, this is generally considered to be the most unequivocal role of *Wolbachia* in host speciation, as (i) there is a growing consensus that speciation rarely occurs due to a single form of reproductive barrier but rather through the accumulation of several (incomplete) isolation barriers (Coyne and Orr, 1997; Presgraves, 2002), and (ii) as unidirectional CI is presumably much more frequent than bidirectional CI.

Third, coevolution of *Wolbachia* and host can accelerate host genetic substitutions and indirectly increase nuclear divergence of host populations. Note that this scenario does not require *Wolbachia* to induce CI; it is rather a general model of

Uninfected population

Resident	Migrant	Hybrid offspring
♀	♂^{W}	partial F_1 inviability, but surviving F_1 (and all subsequent F_i) hybrids are fully viable and fertile
♂	♀^{W}	F_1 hybrids fully viable, but all F_i males in the maternal line are "CI-sterile"

Infected population

Resident	Migrant	Hybrid offspring
♀^{W}	♂	all F_i hybrids fully viable and fertile
♂^{W}	♀	partial F_1 inviability, but surviving F_1 males are fully viable and fertile whereas surviving F_1 females (and their F_i daughters) again suffer partial inviability of their offspring

Table 2.5: Cytoplasmic incompatibility (CI) as a reproductive isolation barrier. CI imposes a reproductive isolation barrier between *Wolbachia*-infected and uninfected host populations. The effect on viability and fertility of hybrids, however, depends on the infection status of the population in which the hybrids are in (i.e., on the environment). Furthermore, effects are sex-dependent. For instance, hybrids that derive from infected male migrants in an uninfected population suffer partial inviability but their surviving offspring (and all subsequent generations) share the resident infection status and are thus fully viable and fertile. "CI-sterility" refers to the inability of infected males to properly fertilize uninfected eggs due to *Wolbachia*'s modification of sperm. In this table, perfect transmission of *Wolbachia* but incomplete cytoplasmic incompatibility is assumed.

how endosymbionts might be a factor in host speciation.

Forth and last, the postzygotic isolation imposed by the expression of CI can select for behavioral isolation through reinforcement (cp. the last Section 2.2.2). In the first two modes of CI-assisted speciation, *Wolbachia* directly contribute to reducing gene flow and enhancing reproductive isolation (Werren, 1998), whereas in the other two modes, the involvement of the bacteria is rather indirect. In Chapters 3, we analyze conditions for the second mode, and in Chapter 4, we devise a method to gauge its contribution to total gene flow reduction. In Chapter 5, we focus on the reinforcement model but will also shed some light on interactions with hybrid incompatibilities in Chapter 6.

In the literature, empirical evidence has been accumulating that the CI-assisted speciation scenarios listed above are not purely hypothetical ones (see, e.g., Table 1 in Engelstädter and Telschow, 2009, containing remarks relating to infection polymorphism and host population structure). Some exemplary findings in which CI indeed imposes a reproductive isolation barrier either between populations of the same species or between closely related sibling species are presented in the following.

Species in which some populations are infected with CI-inducing *Wolbachia* and

other populations uninfected include *Drosophila melanogaster* (Hoffmann et al., 1994), the bug *Laodelphax striatellus* (Hoshizaki and Shimada, 1995), the butterfly *Hypolimnas bolina* (Charlat et al., 2006), and the mite *Tetranychus urticae* (Breeuwer and Jacobs, 1996). Examples of host species polymorphic for *Wolbachia* infections with bidirectional CI are the mosquito *Culex pipiens* (Laven, 1951; Guillemaud et al., 1997; Duron et al., 2006a), *Drosophila simulans* (O'Neill and Karr, 1990; Merçot and Charlat, 2004; James and Ballard, 2000), and the cherry fruit fly *Rhagoletis cerasi* (Riegler and Stauffer, 2002).

Closely related sibling species with different infection statuses and unidirectional CI are for instance the *Drosophila* flies *D. recens* and *D. subquinaria* (Shoemaker et al., 1999), the *Diabrotica* beetles *D. virgifera virgifera* and *D. virgifera zeae* (Giordano et al., 1997), and possibly the two cryptic species of *Eurema* butterflies, *E. hecabe* (Y) and *E. hecabe* (B) (Narita et al., 2006). Furthermore, an example for sibling species infected with different *Wolbachia* strains inducing bidirectional CI is the *Nasonia* wasp species complex, consisting of *N. vitripennis*, *N. giraulti*, and *N. longicornis* (Breeuwer and Werren, 1990).

In the *Gryllus* cricket species complex, Giordano et al. (1997) found complex incompatibility patterns involving both uni- and bidirectional CI. For instance, *G. pennsylvanicus* and *G. firmus* are unidirectionally incompatible (but see Maroja et al., 2008), whereas *G. rubens* and *G. integer* are bidirectionally incompatible. In the interspecies examples, nuclear incompatibilities add to reproductive isolation. Furthermore, patterns of sexual isolation consistent with reinforcement as a response to a CI barrier have been reported in the case of *D. recens* and *D. subquinaria* (Jaenike et al., 2006). We will investigate this case in more detail in Chapter 6.

2.2.4 Models of speciation through sexual selection

Kirkpatrick and Ravigné (2002) reviewed the literature of mathematical models of speciation through the evolution of prezygotic isolation (i.e., by selection and not genetic drift). They identified five major elements that determine the outcome of speciation caused by selection (see Tab. 2.6): the form of disruptive selection, the kind of prezygotic isolating mechanism, how the force of selection is transmitted to the isolating mechanism, the genetic basis of the isolating mechanism, and the initial condition.

One conclusion from this review study is that the combination of direct selection and a one-allele isolating mechanism is much more powerful in favoring speciation than the classical reinforcement scenario, with indirect selection driving a two-allele mechanism. Yet, two-allele mechanisms seem to be more widespread in nature, possibly because the appropriate genetic variation for one-allele mechanisms is relatively rare (Felsenstein, 1981). Kirkpatrick and Ravigné (2002) also argued that geographical isolation —if considered as a genetic locus— can be viewed as an extreme form of assortative mating: only those individuals that carry the same allele at the geography locus are "allowed" to mate (see class II.B in Tab. 2.6). This is in concordance with the claim of Mayr (e.g., 1963) that allopatry is critical in the great majority of speciation events.

The model we will be using in this thesis (see Chapters 5 and 6 for more details) involves elements of the following types:

- I.A: Divergent selection at a trait locus yields different fitnesses in different populations.
- II.A: Prezygotic isolation can evolve through female mating preferences that use a male trait as a cue (the one under disruptive selection).
- II.C(ii): The setting is parapatric with populations linked by migration
- III.B: Selection on female mating preferences is indirect due to cytoplasmically incompatible populations.
- IV.B: There are two alleles at the gene locus for female mating preference.
- V.B: Postzygotic isolation is initially (i.e., after secondary contact) large between populations (due to disruptive selection and cytoplasmic incompatibility).

I. A form of disruptive selection

- A. Fitnesses vary in space
- B. Frequency-dependent natural selection: (i) two niches with independent density regulation, (ii) tension zone, or (iii) competition within a niche
- C. Sexual selection

II. A prezygotic isolating mechanism

- A. Mating preferences
- B. Assortment traits (including geographical isolation)
- C. A geographical setting: (i) allopatry, (ii) parapatry, or (iii) sympatry

III. Transmitting the force of selection to the isolating mechanism

- A. Direct selection
- B. Indirect selection

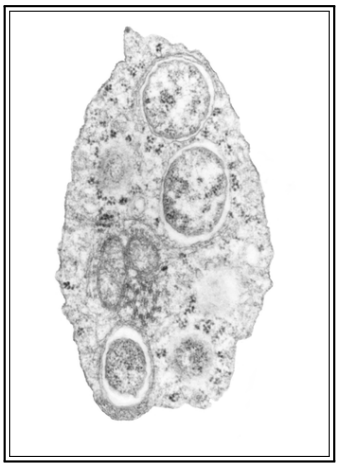
IV. A genetic basis for increased isolation

- A. One-allele mechanisms
- B. Two-allele mechanisms

V. An initial condition

- A. Divergence initially low
- B. Divergence initially large (including geographical isolation)

Table 2.6: Classification of models of speciation through sexual selection. The scheme presented here is taken from Kirkpatrick and Ravigné (2002) who classified 62 models of speciation published between 1966 and 2000 based on five major elements. The model used in this thesis can be classified as I.A / II.A / II.C(ii) / III.B / IV.B / V.B.



Chapter 3

Stability of *Wolbachia* infection polymorphisms

If populations of hosts infected with *Wolbachia* that induce cytoplasmic incompatibility (CI) occur parapatrically with uninfected populations, CI acts as a postzygotic isolation barrier. We investigate the stability of such infection polymorphisms in a mathematical model with two populations linked by migration. We determine critical migration rates below which infected and uninfected populations can coexist. Analytical solutions of the critical migration rate are presented for mainland-island models. These serve as lower estimations for a more general model with two-way migration. The critical migration rate is positive if either *Wolbachia* causes a fecundity reduction in infected females or its transmission is imperfect. We discuss our results with respect to local adaptations of the *Wolbachia* host, speciation, and pest control.

Part of the work presented in this chapter —the special cases *Perfect transmission* and *No fecundity costs* of Section 3.2.5— has been published in the *Journal of Evolutionary Biology* (Flor et al., 2007). The results shown in Section 3.2.6 – *Local adaptation* are part of a publication in *PLoS ONE* (Telschow et al., 2007).

3.1 Introduction

Basic mathematical models of CI show that *Wolbachia* spreads in a single panmictic host population if *Wolbachia* transmission is perfect and the infection does not affect host fecundity (Caspari and Watson, 1959; Fine, 1978). However, if either the transmission is imperfect or the *Wolbachia* infection reduces host fecundity, there exists a threshold infection frequency for the spread of *Wolbachia* (Fine, 1978; Hoffmann et al., 1990). Crucial to understanding these effects is that selection on CI is frequency dependent. Frequency dependence also explains the stability of bidirectional CI (where matings between partners bearing infections with two different *Wolbachia* strains result in fewer offspring) between parapatric host populations (Telschow et al., 2005b) and the spatial spread of *Wolbachia* in host populations due to unidirectional CI (Turelli and Hoffmann, 1991; Schofield, 2002).

A growing number of experimental studies examine the spatial distribution of *Wolbachia*. Although some studies clearly show a spatial spread of CI-inducing *Wolbachia* (Turelli and Hoffmann, 1991; Riegler and Stauffer, 2002; Hiroki et al., 2005), other studies suggest a stable pattern with some populations infected but others not (Shoemaker et al., 1999; Keller et al., 2004; Riegler and Stauffer, 2002; Vala et al., 2004; Jaenike et al., 2006). These studies raise the question under which conditions infected and uninfected host populations can persist in face of migration. This question is especially important because in this case, unidirectional CI acts as a postzygotic isolation barrier between the populations which might be a factor in host speciation as experimental and theoretical work suggest (Shoemaker et al., 1999; Jaenike et al., 2006; Telschow et al., 2007).

In this chapter, we extend the single population model of Fine (1978) and analyse the *Wolbachia* dynamics in two host populations that are linked by migration (see Fig. 3.1). In the focus of our interest are the conditions under which infected and uninfected host populations can stably coexist because only then does unidirectional CI cause postzygotic isolation between the populations. For the purposes of this dissertation, we will use the term infection polymorphism for such a coexistence of infected and uninfected populations. This should not be confused with a polymorphism resulting from infections with different *Wolbachia* strains. We follow Telschow et al. (2005b) and analyse the stability problem in terms of a critical migration rate which is defined as the highest migration rate below which the stable coexistence of infected and uninfected populations is possible. We will demonstrate analytically that the critical migration rate is positive (and hence postzygotic isolation possible) if either *Wolbachia* causes a fecundity reduction in infected females or the transmission rate of *Wolbachia* is imperfect.

Given the existence of an invasion threshold frequency in a panmictic host population (Fine, 1978), it is straightforward to argue that an infection may spread between populations if coupling via migration is strong enough. However, system dynamics and equilibrium states are changed by migration in a non-trivial fashion. Indeed, Nigro and Prout (1990) already conducted computer simulations down to this bottom line. Invasion thresholds derived for an isolated panmictic population cannot be directly applied to a population in the face of migration, and even less to two populations with migration in both directions. In addition, previous theoretical work rather suggests that CI-inducing *Wolbachia* inevitably spread

through spatially structured host populations (Turelli and Hoffmann, 1991; Wade and Stevens, 1994; Schofield, 2002). Hence, based on parameters commonly used to describe *Wolbachia* dynamics, the critical migration rates derived in the present work allow to determine parameter regions of *Wolbachia* spread, of *Wolbachia* extinction, and of stable infection polymorphism. Our results allow to make general conclusions and are therefore a solid base for discussions about infection patterns of *Wolbachia*, its use in pest control and the potential role of unidirectional CI in speciation.

Part of the results presented in this chapter has been published (Flor et al., 2007). Here, we give a generalization of the two special cases presented therein.

3.2 Model and results

We employ recursion equations to model *Wolbachia* dynamics in subsequent host generations. This allows for a stability analysis of *Wolbachia* infection differences between parapatric host populations. In *Section 3.2.6*, we combine these efforts with a simplified population genetic approach to investigate the effect of local adaptations in the host on the stability of infection polymorphisms.

Following Fine (1978), we use three parameters to describe the *Wolbachia* dynamics: level of cytoplasmic incompatibility, l_{CI}; fecundity reduction, f; and transmission rate, t. Population structure is taken into account by the migration rate, m. Tab. 3.1 provides an overview of the parameters and the values they can take. The host organisms reproduce in discrete, non-overlapping generations. Hosts may be either infected with *Wolbachia* or not. *Wolbachia* is assumed to be transmitted strictly maternally via egg cytoplasm. However, not all offspring of infected females may inherit the infection. We define the *Wolbachia* transmission rate, t, as the fraction of infected eggs among all eggs of an infected female. Further, infected female hosts may suffer a fecundity reduction (note that in Fine's model (1978), both sexes' fecundity was assumed to be affected by *Wolbachia*), i.e., the number of eggs is reduced by a certain fraction, f. Cytoplasmic incompatibility is described by the CI level, l_{CI}, which is defined as the proportion of zygotes that die if the egg was uninfected but fertilized by sperm from an infected male. All three parameters may range from zero to one. Experimental studies have shown that the transmission rate is usually very high, at 95 – 100% (see Hoffmann and Turelli, 1997, for a review), whereas fecundity reductions are low or absent, e.g. 10 – 20% in *Drosophila simulans* (Hoffmann et al., 1990). CI levels are much more variable; for example in the *Drosophila* species complex, they range from low levels of less than 10% up to nearly complete incompatibility (Bourtzis et al., 1996; Merçot and Charlat, 2004). Also note that CI levels have been found to be dynamic. For example in *D. melanogaster*, they decline with male age (Reynolds et al., 2003). However, in our model host generations do not overlap. CI levels can therefore be interpreted as an average over the host's life span.

We will analyse the *Wolbachia* dynamics in three different population structures shown in Fig. 3.1: (a) A mainland-island model with an uninfected mainland, (b) a mainland-island model with an infected mainland, and (c) a model with two populations and two-way migration. In all three cases, migration is described by the migration rate, m; i.e., each generation, a fraction m of the target population

Symbol	Description
f	Fecundity costs of a *Wolbachia* infection—infected females' egg production is reduced by f; $0 \leq f < 1$
l_{CI}	Level of cytoplasmic incompatibility (CI level)—fraction of zygotes that die if an uninfected egg is fertilized by sperm from an infected male; $0 \leq l_{\mathrm{CI}} \leq 1$
m	Migration rate—fraction of a host population that is replaced by migrants; $0 \leq m \leq 1$
s	Selection coefficient—hosts carrying a locally adaptive allele at the nuclear locus T have increased viability by factor $1+s$; $s \geq 0$
t	Transmission rate—fraction of an infected females' eggs that inherit *Wolbachia*; $0 < t \leq 1$
x	*Wolbachia* infection frequency within a host population
T_i	Allele at host nuclear trait locus T; $i \in \{1, 2\}$
l_{crit}	Critical CI level—*Wolbachia* may stably persist in a host population if the CI level is above the critical level
m_{crit}	Critical migration rate—infection polymorphism between populations may be stable if migration is below the critical migration rate

Table 3.1: Glossary of model notation. The ranges displayed for model parameters refer to values that are mathematically valid and biologically meaningful in the model context, but not necessarily reasonable. E.g., a migration rate of $m = 1$ would mean that a host population becomes completely replaced by migrants. This is a valid mathematical assumption and can be interpreted biologically as well. However, it is not a reasonable assumption. For biologically reasonable parameter values, see main text.

is replaced by migrants from the source population. In the model with two-way migration, we allow migration rates to be different in both directions. We define m_i as the fraction of population i that is replaced by migrants from the other population. In our model, the migration rate may take values between zero and one.

We start our analysis with a summary of the dynamics of *Wolbachia* within a single panmictic host population. These results will be used as a baseline to understand the infection dynamics of parapatric host populations, i.e., of the mainland-island models and the model with two-way migration. Because we are interested in the stability of infection polymorphism, we always include starting conditions with one population infected but the other not. We first present results for the general model, and then analyse three special cases. In each of these cases, one parameter is restricted to take only one value: (i) $t = 1$ (perfect *Wolbachia* transmission, *Section 3.2.5.1*), (ii) $f = 0$ (no fecundity costs, *Section 3.2.5.2*), and (iii) $l_{\mathrm{CI}} = 1$ (complete incompatibility, *Section 3.2.5.3*). The results for the first two special cases have been published in the *Journal of Evolutionary Biology* (Flor et al., 2007).

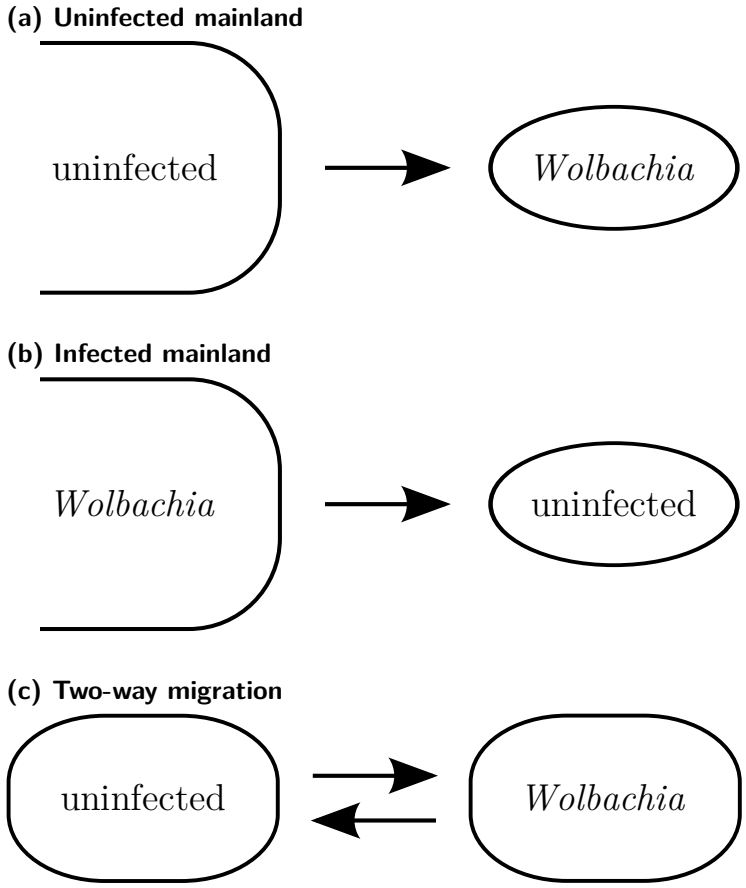


Figure 3.1: Population structure. We investigated three scenarios where unidirectional CI causes postzygotic isolation between parapatric host populations. Subfigures (a) and (b) show mainland-island models with migration occurring only from the mainland to the island. In Subfigure (c), two populations are linked by migration in both directions. Note that the graphs show only starting conditions. After the onset of migration (secondary contact), both infected and uninfected individuals can occur in the same population.

3.2.1 Single host population

Prior to analyzing the stability of infection polymorphism in the context of parapatric host populations, we need to understand the *Wolbachia* dynamics in a single panmictic host population.

Dynamics. We follow Fine (1978) and describe the dynamics of unidirectional CI in a panmictic host population by a recursion equation. Let x and x' denote the *Wolbachia* frequencies in consecutive generations. x' can be calculated by way of recursion from x as the fraction of infected offspring among all offspring. Offspring production depends on the infection states of the mating partners, as shown in Tab. 3.2. E.g. if infected females and males mate with each other then a relative number of $(1-f)\,(1-l_{\mathrm{CI}})\,(1-t)$ uninfected offspring and $(1-f)t$ infected offspring is produced, compared to the offspring from matings where both partners are uninfected. The probability of matings is assumed to follow a mass action law, based on

♀	♂		offspring infection status
	1	$1 - l_{\mathrm{CI}}$	uninfected
	0	0	infected
	$(1-f)(1-t)$	$(1-f)(1-l_{\mathrm{CI}})(1-t)$	uninfected
	$(1-f)\,t$	$(1-f)\,t$	infected

Table 3.2: Mating table. The relative number of offspring produced from different mating pairs depends on the fecundity reduction, f, the CI level, l_{CI}, and the transmission rate, t. E.g. if *Wolbachia (W)* infected females (♀) and males (♂) mate then $(1-f)(1-l_{\mathrm{CI}})(1-t)$ uninfected and $(1-f)\,t$ infected offspring are produced.

the frequencies of infected and uninfected hosts. Thus, considering unidirectional CI, the transmission of *Wolbachia*, and their effects on female fecundity, we can formulate the infection dynamics in the following way:

$$x' = \frac{(1-f)\,t\,x}{1 - fx\big[1 - l_{\mathrm{CI}}\,(1-t)\,x\big] - l_{\mathrm{CI}}\,x\,(1-tx)} \equiv F(f, l_{\mathrm{CI}}, t, x). \tag{3.1}$$

In Fig. 3.2, the infection dynamics are illustrated as an iterative map for $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, and $t = 0.9$.

Fixpoints. Equation (3.1) has three fixpoints which can be calculated by requiring $x' = x \equiv x^*$ and solving for x^*:

$$x_1^* = 0, \tag{3.2a}$$

$$x_{2,3}^* = \frac{f + l_{\mathrm{CI}} \mp \sqrt{R(f, l_{\mathrm{CI}}, t)}}{D(f, l_{\mathrm{CI}}, t)}, \tag{3.2b}$$

where

$$R(f, l_{\mathrm{CI}}, t) = (l_{\mathrm{CI}} - f)^2 - 4\,(1-f)^2\,l_{\mathrm{CI}}\;t\,(1-t)\,, \text{ and} \tag{3.3a}$$

$$D(f, l_{\mathrm{CI}}, t) = 2\,l_{\mathrm{CI}}\big[f + (1-f)\,t\big] \tag{3.3b}$$

are two functions we introduce to allow for a more perspicuous notation of the above fixpoints and future equations. If $f = 0$ and $t = 1$ then $R = l_{\mathrm{CI}}^2$, and $x_1^* = x_2^* = 0$ is unstable, whereas $x_3^* = 1$ is stable. *Wolbachia* goes to fixation for any positive level of CI in that case. However, if either $f > 0$ or $t < 1$ then fixpoint x_1^* is stable for any level of CI and represents an equilibrium where *Wolbachia* is absent from the population. Furthermore, if $f > 0$ or $t < 1$ then *Wolbachia* can persist stably

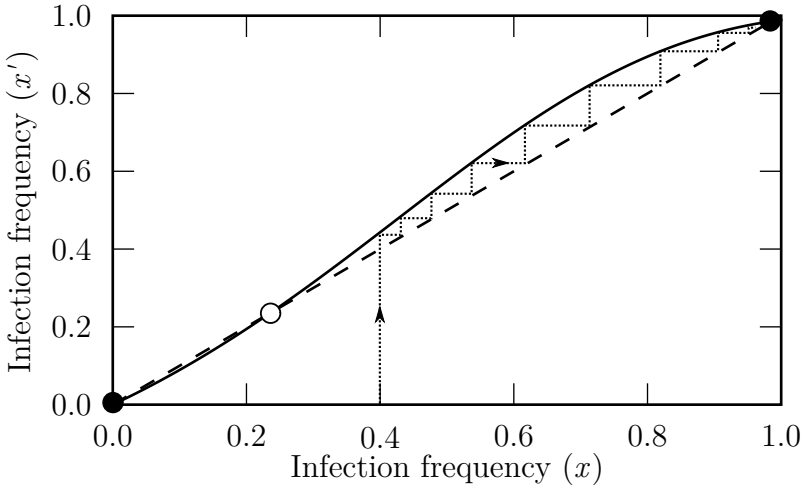


Figure 3.2: *Wolbachia* infection dynamics in a panmictic host population. The graph shows the infection dynamics as an iterative map of Eq. (3.1). Fixpoints are given by intersections with the dashed diagonal (where $x = x'$) and are marked by circles (filled – stable, open – unstable). The dotted line with midarrows illustrates the infection dynamics starting from a frequency of 0.4, which ultimately leads to close fixation of *Wolbachia*. Parameters are: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $t = 0.9$.

in the population if and only if (i) $x^*_{2,3}$ are real numbers and (ii) $0 \leq x^*_2 < 1$. These conditions are fulfilled if and only if

$$l_{\mathrm{CI}} > l_{\mathrm{crit}}(f,t) = f + Q(f,t) + \sqrt{\left[f + Q(f,t)\right]^2 - f^2}, \tag{3.4a}$$

$$t > \tfrac{1}{2}, \tag{3.4b}$$

where $Q(f,t) = 2\,(1-f)^2\,t\,(1-t)$. Inequality (3.4a) defines a critical CI level, $l_{\mathrm{crit}} = l_{\mathrm{crit}}(f,t)$, and guarantees that $x^*_{2,3}$ are real numbers. If this is satisfied, and additionally Inequality (3.4b) holds then it is ensured that $0 \leq x^*_2 < 1$. Note that this implies $0 < x^*_3 \leq 1$. If both conditions are met then x^*_3 is stable and characterizes a state where *Wolbachia* has reached a stable equilibrium frequency. Furthermore, x^*_2 is unstable and marks a frequency threshold determining whether the system converges towards x^*_1 (if starting from below x^*_2) or x^*_3 (if starting from above x^*_2). In Fig. 3.2, Inequalities (3.4) are both satisfied.

Inequality (3.4b) states that $t = \frac{1}{2}$ represents a lower boundary for the transmission rate in terms of *Wolbachia* persistence in a population. Note that for perfect *Wolbachia* transmission, $t = 1$, the critical CI level simplifies to $l_{\mathrm{crit}}(f,1) = f$.

The most important insights from the analysis in this section are that a *Wolbachia* infection can stably persist in a panmictic population only if the CI level is above a critical threshold, and that a transmission rate of $t = \frac{1}{2}$ represents a lower boundary for such a stable persistence (compare Fine, 1978; Hoffmann et al., 1990). We will refer to these results repeatedly below.

3.2.2 Uninfected mainland

Building upon the dynamics within a single panmictic host population, we discuss the effects of migration on the dynamics and persistence of *Wolbachia* (compare Fig. 3.1). In the next two sections, we allow migration to occur only in one direction, from a mainland to an island population. We will consider two different scenarios. First, we analyse the case of an uninfected mainland (this section). Second, we will assume that the *Wolbachia* infection is at fixation in the mainland population (see Section 3.2.3). In a third and most general scenario, we allow migration to occur in both directions (Section 3.2.4).

Dynamics. We analyse the dynamics of *Wolbachia* in an island population receiving migration from an uninfected mainland (see Fig. 3.1a). Let x and x' denote the *Wolbachia* frequencies on the island in consecutive generations. Employing function $F(f, l_{\mathrm{CI}}, t, x)$, defined in Eq. (3.1), we can state the following difference equation describing the *Wolbachia* dynamics on the island:

$$x' = (1-m)\,F(f, l_{\mathrm{CI}}, t, x) \equiv G(f, l_{\mathrm{CI}}, m, t, x). \tag{3.5}$$

Fixpoints. In order to determine the fixpoints x^* of this system, we solve Eq. (3.5) for $x^* = x' = x$. This yields

$$x_1^* = 0, \tag{3.6a}$$

$$x_{2,3}^* = \frac{f + l_{\mathrm{CI}} \mp \sqrt{(l_{\mathrm{CI}} - f)^2 - 4\,(1-f) l_{\mathrm{CI}} t \big[m + (1-f)(1-m)(1-t)\big]}}{2\,l_{\mathrm{CI}}\big[f + (1-f)t\big]} \tag{3.6b}$$

$$= \frac{f + l_{\mathrm{CI}} \mp \sqrt{\hat{R}(f, l_{\mathrm{CI}}, m, t)}}{D(f, l_{\mathrm{CI}}, t)},$$

where

$$\hat{R}(f, l_{\mathrm{CI}}, m, t) = (l_{\mathrm{CI}} - f)^2 - 4\,(1-f)\,l_{\mathrm{CI}}\,t\,\big[m + (1-f)\,(1-m)\,(1-t)\big] \tag{3.7}$$

is an extension of $R(f, l_{\mathrm{CI}}, t)$ (compare Eq. (3.3a)) for the scenario with an uninfected mainland. For $m = 0$ it holds that $\hat{R}(f, l_{\mathrm{CI}}, 0, t) = R(f, l_{\mathrm{CI}}, t)$. Thus, without migration Equations (3.6) correctly yield the fixpoints for an isolated host population (compare Equations (3.2)).

A stability analysis shows that, if $m > 0$ then x_1^* is stable for all allowed values of the other parameters: $0 \le f < 1$, $0 \le l_{\mathrm{CI}} \le 1$, and $0 < t \le 1$ (compare Tab. 3.1). *Wolbachia* can persist on the island if and only if (i) $x_{2,3}^*$ are real numbers, and if additionally (ii) $0 \le x_2^* < 1$. Otherwise, x_1^* is the only stable equilibrium frequency, and *Wolbachia* goes to extinction on the island for arbitrary starting conditions. If both conditions are met, it holds that $0 < x_3^* < 1$; and x_3^* is stable and describes the situation where *Wolbachia* can persist on the island despite permanent migration of uninfected individuals from the mainland. Moreover, fixpoint x_2^* is unstable in that case and denotes a threshold frequency (unstable equilibrium frequency). If the frequency of *Wolbachia* on the island is above this threshold at the beginning then the system converges towards x_3^*; but if it is below x_2^*, it converges towards

x_1^*. Conditions (i) and (ii) are fulfilled if and only if

$$l_{\mathrm{CI}} > \hat{l}_{\mathrm{crit}}(f,m,t) = f + \hat{Q}(f,m,t) + \sqrt{\left[f + \hat{Q}(f,m,t)\right]^2 - f^2}, \tag{3.8a}$$

$$t > \tfrac{1}{2}, \tag{3.8b}$$

where $\hat{Q}(f,m,t) = 2\,(1-f)\left[(1-f\,(1-m))\;t - (1-f)\,(1-m)\,t^2\right]$. Because it holds that $\hat{Q}(f,0,t) = Q(f,t)$ and consequently $\hat{l}_{\mathrm{crit}}(f,0,t) = l_{\mathrm{crit}}(f,t)$, Inequality (3.8a) simplifies to Inequality (3.4a) if we assume $m = 0$. The critical CI level increases strictly monotonously with higher migration rates. This reflects the requirement for *Wolbachia* to induce ever stronger incompatibility in order to avoid extinction in the face of migration of uninfected hosts. In the next section, we follow Telschow et al. (2005b) and describe the stability of the infection polymorphism in terms of a critical migration rate rather than a critical CI level.

Critical migration rate. For the model with an uninfected mainland, Telschow et al. (2005b) defined the critical migration rate as the highest migration rate below which *Wolbachia* can stably persist on the island. The existence of the critical migration rate follows from the strictly monotone decrease of the square root in Eq. (3.6b) if interpreted as a function of m, and the fact that this function has a root. Figure 3.3a illustrates the critical migration rate by showing the equilibrium infection frequencies as a function of the migration rate. For low migration rates, three equilibrium frequencies exist. With increasing migration, the equilibrium frequency x_3^* decreases while the threshold frequency x_2^* increases. The distance between the two frequencies becomes smaller until they coincide at the critical migration rate. For migration rates above this critical value, the infection frequency converges towards x_1^* for arbitrary starting frequencies.

The critical migration rate can be calculated analytically. We first note that the critical migration rate must be a non-negative number. This can be understood by considering the case that $l_{\mathrm{CI}} \leq l_{\mathrm{crit}}$. Under these circumstances, the infection cannot stably persist in the island population even without migration of uninfected hosts, as shown in Section 3.2.1. Hence, the critical migration rate is zero in this case. However, for $l_{\mathrm{CI}} > l_{\mathrm{crit}}$, positive critical migration rates occur. As Telschow et al. (2005b) have pointed out, the critical migration rate, denoted here by $m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}$ (superscript u→w denotes migration from an uninfected mainland to a *Wolbachia* infected island), is the migration rate for which $x_2^* = x_3^*$. From Equations (3.6b) and (3.7), it follows that the critical migration rate must fulfill the equation $\hat{R}(f, m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}, t) = 0$, i.e.

$$(l_{\mathrm{CI}} - f)^2 - 4\,(1-f)\,l_{\mathrm{CI}}\,m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}\,t = 4\,(1-f)^2\,l_{\mathrm{CI}}\,(1-m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}})\,t\,(1-t)\,. \tag{3.9}$$

The critical migration rate for the scenario with an uninfected mainland can therefore be written as a function of the fecundity reduction, the CI level, and the

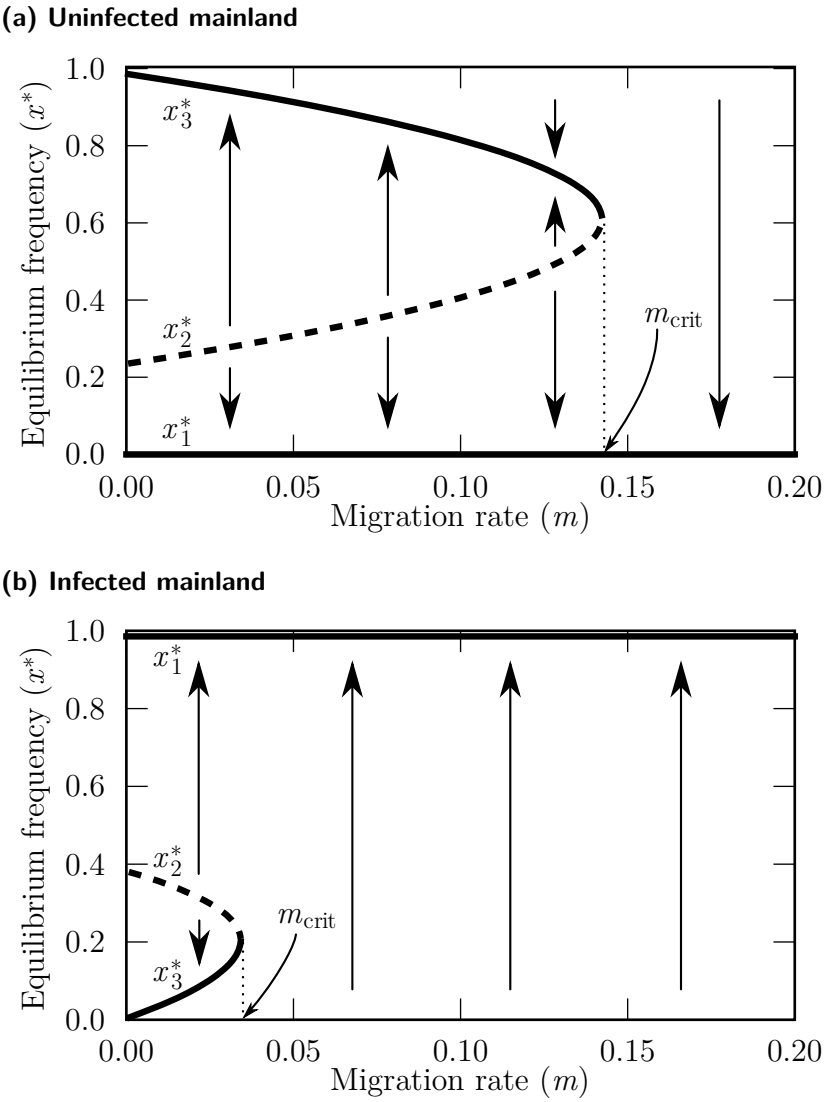


Figure 3.3: *Wolbachia* equilibrium frequencies for mainland-island models. Shown are the real fixpoints for (a) the model with an uninfected mainland and (b) the model with an infected mainland. Stable fixpoints are denoted by solid, unstable ones by dashed lines. The vertical arrows indicate trajectories of the system for fixed migration rates. The critical migration rate, m_{crit}, is pointed up by dotted lines and in both graphs marks the (saddle-node) bifurcation where the fixpoints x_2^* and x_3^* coincide. For migration rates above the critical migration rate, x_1^* is the only real fixpoint, and the system converges towards it for any starting condition. A stable polymorphic equilibrium is only possible if $m < m_{\text{crit}}$. Note that the infection polymorphism is considerably more stable if the mainland is uninfected rather than infected. Parameters: $f = 0.1$, $l_{\text{CI}} = 0.9$, $t = 0.9$.

transmission rate:

$$\begin{aligned} m_{\text{crit}}^{\text{u}\to\text{w}}(f, l_{\text{CI}}, t) &= \frac{(l_{\text{CI}} - f)^2 - 4(1-f)^2 l_{\text{CI}}\, t\,(1-t)}{4(1-f)\, l_{\text{CI}}\, t\,[f + (1-f)\, t]} \\ &= \frac{R(f, l_{\text{CI}}, t)}{2(1-f)\, t\, D(f, l_{\text{CI}}, t)} \end{aligned} \tag{3.10}$$

if $l_{\text{CI}} > l_{\text{crit}}$. If $l_{\text{CI}} \leq l_{\text{crit}}$, the critical migration rate is zero. Note that the inequality $m < m_{\text{crit}}^{\text{u}\to\text{w}}$ is equivalent to Inequality (3.8a). Together with the requirement that $t > \frac{1}{2}$, both guarantee that a stable infection polymorphism between an uninfected mainland and a mostly infected island is possible.

Figures 3.4a and 3.5 show how the critical migration rate depends on the three parameters, f, l_{CI}, and t. Fig. 3.4a illustrates that the critical migration rate is zero for $l_{\text{CI}} \leq l_{\text{crit}}$. For higher levels of CI, however, it increases strictly monotonously. This is because, for increasing levels of CI, an increasing number of uninfected migrants is needed to supplant the infection on the island. The same line of thought explains why the critical migration rate increases when *Wolbachia* transmission becomes more efficient (see Fig. 3.5b). Figure 3.5a on the other hand shows that if CI level and transmission rate are fixed then the critical migration rate decreases strictly monotonously with increasing fecundity reduction. In this case, less migration is needed to supersede the infection that is now linked to growing fitness costs. Note that the highest possible critical migration rate for this scenario is $m_{\text{crit}}^{\text{u}\to\text{w}} = \frac{1}{4}$. It is achieved for $f = 0$, $l_{\text{CI}} = 1$, and $t = 1$.

3.2.3 Infected mainland

Next, we analyse the situation where an island receives migration from a mainland with a stable *Wolbachia* infection.

Dynamics. We assume that in the mainland population, the infection is at the stable equilibrium frequency

$$x_{\text{main}}(f, l_{\text{CI}}, t) = \frac{f + l_{\text{CI}} + \sqrt{R(f, l_{\text{CI}}, t)}}{D(f, l_{\text{CI}}, t)}, \tag{3.11}$$

as derived for an isolated host population (see Eq. (3.2b)). Then, we can use function $F(f, l_{\text{CI}}, t, x)$, defined in Eq. (3.1), to state a difference equation describing the *Wolbachia* dynamics on the island:

$$x' = (1-m)\, F(f, l_{\text{CI}}, t, x) + m\, x_{\text{main}}(f, l_{\text{CI}}, t) \equiv H(f, l_{\text{CI}}, m, t, x). \tag{3.12}$$

Note that these dynamics imply that conditions (3.4) are satisfied, i.e., $l_{\text{CI}} > l_{\text{crit}}$ and $t > \frac{1}{2}$, because otherwise the infection could not persist on the mainland, and the additive term "$+\, m\, x_{\text{main}}$" would not be appropriate.

Fixpoints. First note that $x_{\text{main}}(f, l_{\text{CI}}, t)$ must be a fixpoint of Eq. (3.12). Because $H(f, l_{\text{CI}}, m, t, x)$ is a polynomial in x of degree 3, a degree 2 polynomial can be

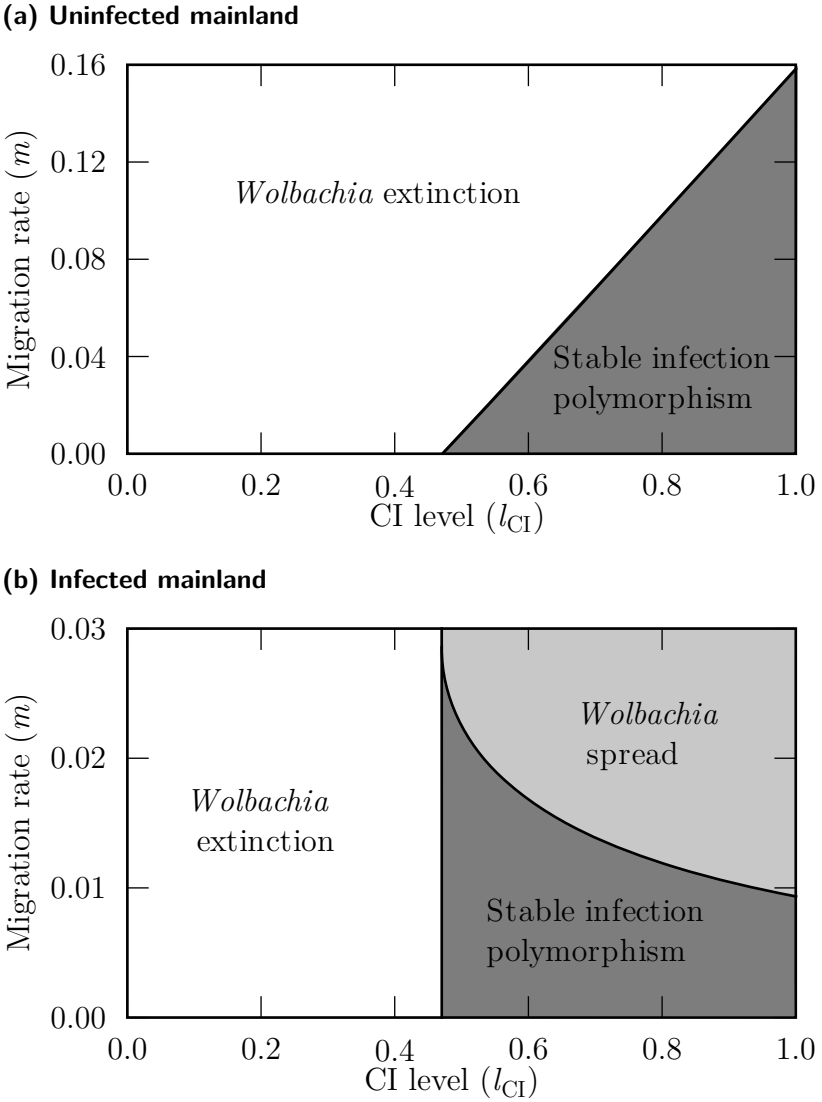


Figure 3.4: Critical migration rates for the mainland-island models. The critical migration rate, m_{crit}, divides the parameter plane spanned by CI level and migration rate into regions with different infection patterns. If $l_{\text{CI}} > l_{\text{crit}}(f,t) \approx 0.47$ and $m < m_{\text{crit}}$ then the infection polymorphism is stable (dark gray region). For all other parameter constellations, the polymorphism is destroyed, either by *Wolbachia* spread (light gray region) or by *Wolbachia* extinction (white region). Parameters: $f = 0.1$, $t = 0.9$.

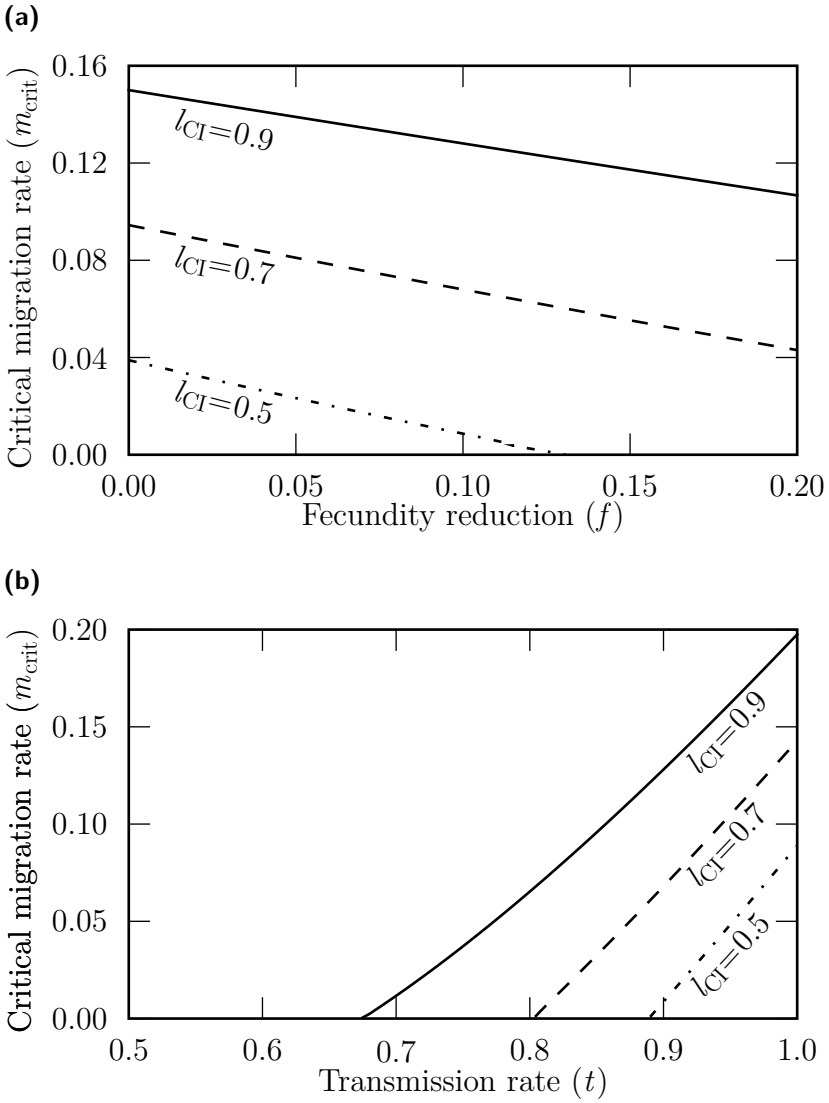


Figure 3.5: Critical migration rates for the mainland-island model with an uninfected mainland. If migration is below the critical migration rate, the infection polymorphism is stable. If migration is larger then the infection polymorphism is destroyed as *Wolbachia* spreads onto the island population. For a given level of CI, the critical migration rate (a) decreases with increasing fecundity reduction, f, and (b) increases with increasing transmission rate, t. Parameters: (a) $t = 0.9$; (b) $f = 0.1$.

calculated by polynomial division:

$$\hat{H}(f, l_{\mathrm{CI}}, m, t, x) = \frac{H(f, l_{\mathrm{CI}}, m, t, x) - x}{x - x_{\mathrm{main}}(f, l_{\mathrm{CI}}, t)}. \tag{3.13}$$

Eq. (3.13) has two roots which are fixpoints of the original dynamical system, $x' = H(f, l_{\mathrm{CI}}, m, t, x)$, and which can be found analytically. Consequently, the three fixpoints of Eq. (3.12) are

$$x_1^* = \frac{f + l_{\mathrm{CI}} + \sqrt{R(f, l_{\mathrm{CI}}, t)}}{D(f, l_{\mathrm{CI}}, t)} = x_{\mathrm{main}}(f, l_{\mathrm{CI}}, t), \tag{3.14a}$$

$$\begin{aligned} x_{2,3}^* = &\frac{(f + l_{\mathrm{CI}})(1 + m) - (1 - m)\sqrt{R(f, l_{\mathrm{CI}}, t)}}{2\, D(f, l_{\mathrm{CI}}, t)} \\ &\pm \frac{\sqrt{\left[(f + l_{\mathrm{CI}})(1 + m) - (1 - m)\sqrt{R(f, l_{\mathrm{CI}}, t)}\right]^2 - 8\, m\, D(f, l_{\mathrm{CI}}, t)}}{2\, D(f, l_{\mathrm{CI}}, t)} \end{aligned} \tag{3.14b}$$

The functions $R(f, l_{\mathrm{CI}}, t)$ and $D(f, l_{\mathrm{CI}}, t)$ were introduced in Equations (3.3). For $m = 0$, the fixpoints of a single host population are correctly reproduced by Equations (3.14), since $x_2^*|_{m=0} = \frac{f + l_{\mathrm{CI}} - \sqrt{R(f, l_{\mathrm{CI}}, t)}}{D(f, l_{\mathrm{CI}}, t)}$ and $x_3^*|_{m=0} = 0$. Note however that the fixpoint indexes are swapped compared to the single population scenario because x_1^* in the infected mainland scenario corresponds to x_3^* in a single population. In Fig. 3.3b, it is shown how the three fixpoint frequencies depend on the migration rate.

If $f \geq 0$ and $m > 0$ then x_1^* is stable for all parameter combinations that satisfy Inequalities (3.4), and describes the state of *Wolbachia* spread onto the island. A stable infection polymorphism is maintained if the spread of *Wolbachia* is prevented. This is possible if and only if $x_{2,3}^*$ are real numbers. Then, x_3^* is stable whereas x_2^* is unstable. x_3^* characterizes a state of low infection frequency due to recurrent migration of infected individuals, and x_2^* again marks a threshold frequency.

A critical CI level is not a proper measure for the stability of an infection polymorphism in the infected mainland scenario. This is because we are interested in the existence of the fixpoint which represents a state of *low* infection frequency. Inequalities (3.4) only ensure that the fixpoint of *high* prevalence of *Wolbachia* exists. Migration of mainly infected hosts from the mainland population benefits *Wolbachia* on the island; thus, it would in principal lower the level of incompatibility necessary for the spread of *Wolbachia*. However, lowering the level of CI below the critical level drives the infection to extinction on the mainland. In conclusion, for the scenario with an infected mainland population the critical migration rate is the appropriate measure of infection polymorphism stability as we will show in the next section.

Critical migration rate. According to our analysis, a stable infection polymorphism —i.e., an infected mainland and a (mainly) uninfected island— is possible if and only if the island infection frequency is at x_3^*. Again, the stability can be described by a critical migration rate, denoted by $m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}$ (superscript w→u designates migration from a *Wolbachia* infected mainland to an uninfected island). Here, the critical migration rate is defined as the highest migration rate below which the island

population can maintain a state of low *Wolbachia* frequency. Both x_2^* and x_3^* are real numbers if $m < m_{\text{crit}}^{\text{w}\to\text{u}}$. For migration rates above this critical value, the infection converges towards x_1^* independent of its starting frequency. For $l_{\text{CI}} \leq l_{\text{crit}}$, the critical migration rate is zero because the infection does not persist in the mainland population (compare Section 3.2.1). However, for $l_{\text{CI}} > l_{\text{crit}}$, the critical migration rate is positive and can be computed by requiring that $x_2^* = x_3^*$. Combining these considerations, the critical migration rate for the infected mainland scenario is

$$\begin{aligned} m_{\text{crit}}^{\text{w}\to\text{u}}(f, l_{\text{CI}}, t) &= 4\, l_{\text{CI}} \big[f + (1-f)\, t\big] \\ &\quad \cdot \left(\frac{1 - \sqrt{(1-f)t}}{f + l_{\text{CI}} + \sqrt{(l_{\text{CI}} - f)^2 - 4\,(1-f)^2\, l_{\text{CI}}\; t\,(1-t)}} \right)^2 \\ &= 2\, D(f, l_{\text{CI}}, t) \left(\frac{1 - \sqrt{(1-f)t}}{f + l_{\text{CI}} + \sqrt{R(f, l_{\text{CI}}, t)}} \right)^2 \end{aligned} \tag{3.15}$$

if $l_{\text{CI}} > l_{\text{crit}}$. If $l_{\text{CI}} \leq l_{\text{crit}}$, the critical migration rate is zero.

In Fig. 3.4b, the critical migration rate is plotted as a function of the CI level. As can be seen, the infection polymorphism is stable if $l_{\text{CI}} > l_{\text{crit}}$, and if migration is below the critical rate. For higher migration rates, the *Wolbachia* infection spreads to the island and gets close to fixation (only hampered by imperfect transmission) in both populations, hence destroying the infection polymorphism. For fixed fecundity reduction and transmission rate, the potential of *Wolbachia* to spread becomes greater with increasing CI levels, resulting in decreasing critical migration rates. Due to the same reason, increasing fecundity reductions and decreasing transmission rates lead to higher critical migration rates for constant CI levels (Fig. 3.6). The highest critical migration rates are achieved if CI levels are barely larger than l_{crit}. Then, the effect of CI is merely strong enough to keep *Wolbachia* within the mainland population, and —especially if fecundity reductions are large and/or transmission rates are low— high migration rates are necessary to permit the spread onto the island.

3.2.4 Two-way migration

The model can be generalized to incorporate migration between two populations in both directions (see Fig. 3.1c). In this section, we model both symmetric and asymmetric migration between the populations.

Dynamics. After reproduction has occurred, a fraction m of each population is replaced by migrants from the other one. Let the populations be labeled A and B from left to right as depicted in Fig. 3.1c. Then, denoting with x_{A} and x_{B} the infection frequencies in populations A and B, respectively, the system dynamics become a set of two coupled difference equations:

$$x'_{\text{A}} = (1 - m_{\text{A}})\; F(f, l_{\text{CI}}, t, x_{\text{A}}) + m_{\text{A}} x_{\text{B}}, \tag{3.16a}$$

$$x'_{\text{B}} = (1 - m_{\text{B}})\; F(f, l_{\text{CI}}, t, x_{\text{B}}) + m_{\text{B}} x_{\text{A}}, \tag{3.16b}$$

where the function $F(f, l_{\mathrm{CI}}, t, x)$ models reproduction within each population according to Eq. (3.1).

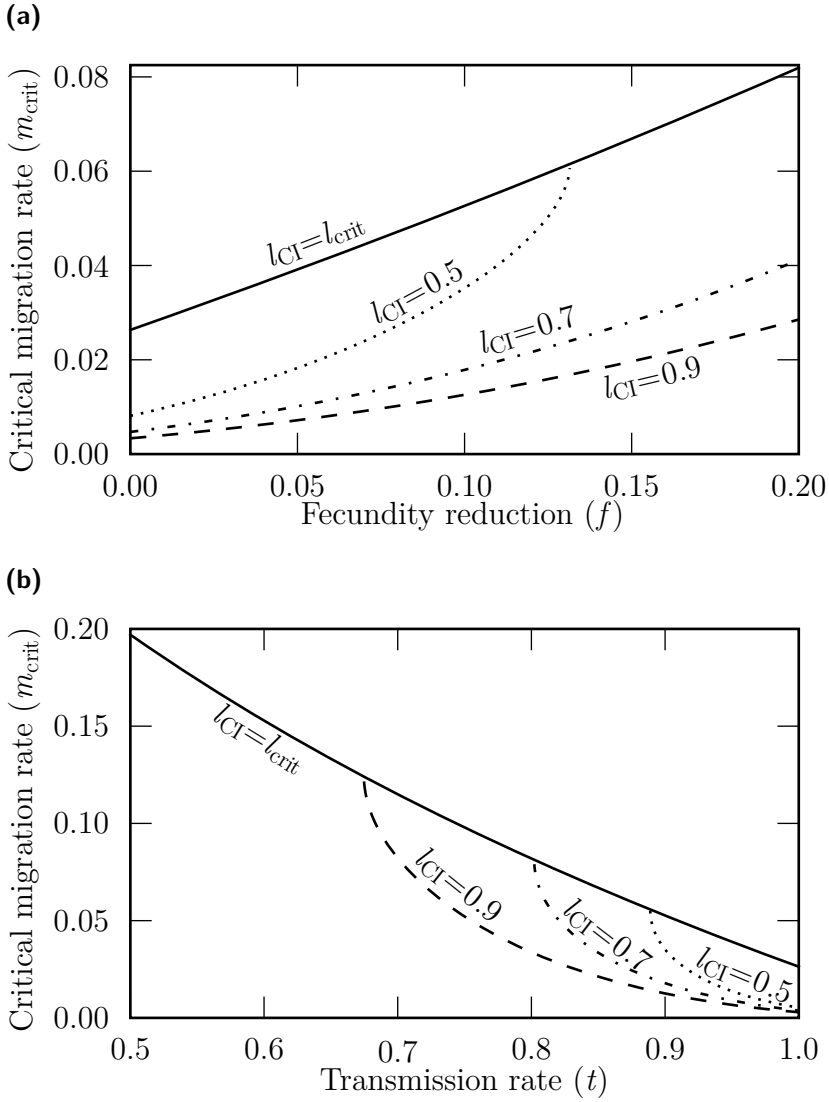


Figure 3.6: Critical migration rates for the mainland-island model where the mainland is infected with *Wolbachia*. The infection polymorphism can only be stable if migration is below the critical migration rate; and it is inevitably destroyed if migration is larger because then *Wolbachia* spreads onto the island. (a) For a given level of CI, the critical migration rate increases with increasing fecundity reduction, f. For fecundity reductions beyond a critical value —i.e., when the considered CI level equals the threshold $l_{\mathrm{crit}}(f, t)$— the infection becomes lost in both populations. (b) Similarly, the critical migration rate decreases with increasing transmission rate, t, and *Wolbachia* goes to extinction if its transmission is below a critical rate. In both cases, the line labeled $l_{\mathrm{CI}} = l_{\mathrm{crit}}$ links all those threshold points and thus marks the upper limit of the critical migration rate. Parameters: (a) $t = 0.9$; (b) $f = 0.1$.

Numerical simulations and equilibrium states. In contrast to the models with one-way migration, we were not able to solve the bidirectional migration system analytically. Therefore, we conducted numerical simulations to determine equilibrium states for the two populations. We started all of our simulations with population A being completely uninfected and population B consisting exclusively of infected hosts (compare Fig. 3.1c). We numerically iterated the dynamics for 10^6 generations or until an equilibrium had been reached. We considered the system to be in equilibrium if *Wolbachia* frequency differences between two consecutive generations were smaller than 10^{-11} and still declining.

We screened the parameter space of the two-way migration model in order to determine under which conditions a stable infection polymorphism is possible. In general, the simulations confirmed that *Wolbachia* cannot stably persist if $l_{\mathrm{CI}} \leq l_{\mathrm{crit}}$. This is true for arbitrary migration rates. In addition, we found that a stable coexistence of an infected and an uninfected host population is possible if (i) $l_{\mathrm{CI}} > l_{\mathrm{crit}}$ and if (ii) migration rates stay within certain parameter regions. Figure 3.7 illustrates these findings for symmetric migration, $m = m_{\mathrm{A}} = m_{\mathrm{B}}$. Fecundity reduction and *Wolbachia* transmission are fixed at $f = 0.1$ and $t = 0.9$, respectively. In the parameter plane spanned by migration rate and CI level, three regions can be distinguished: (i) a region where a stable infection polymorphism is possible, (ii) a region where the *Wolbachia* infection spreads to the same equilibrium frequency in both populations, and (iii) a region where the infection is lost in both populations. Denoting the equilibrium frequencies found in the two populations by $x_{\mathrm{A}}^{(\mathrm{eq})}$ and $x_{\mathrm{B}}^{(\mathrm{eq})}$, these three regions can be described in the following manner:

$$\text{(i)} \quad x_{\mathrm{A}}^{(\mathrm{eq})} < x_{\mathrm{B}}^{(\mathrm{eq})}, \tag{3.17a}$$

$$\text{(ii)} \quad x_{\mathrm{A}}^{(\mathrm{eq})} = x_{\mathrm{B}}^{(\mathrm{eq})} = x_3^*(f, l_{\mathrm{CI}}, t) \quad \text{by Eq. (3.2b)}, \tag{3.17b}$$

$$\text{(iii)} \quad x_{\mathrm{A}}^{(\mathrm{eq})} = x_{\mathrm{B}}^{(\mathrm{eq})} = x_1^* = 0 \qquad \text{by Eq. (3.2a)}. \tag{3.17c}$$

In the latter two cases, infection polymorphism and hence reproductive isolation between the populations is destroyed.

For asymmetric migration, Fig. 3.8 depicts equilibrium states of the system in the parameter plane spanned by the two migration rates, m_{A} and m_{B} (fecundity reduction, CI level, and transmission rate are fixed at $f = 0.1$, $l_{\mathrm{CI}} = 0.5$, and $t = 0.9$, respectively). Because $l_{\mathrm{CI}} > l_{\mathrm{crit}}$, a region exists where the mainly uninfected population A and the mainly infected population B can stably coexist. However, outside of this region, *Wolbachia* either spreads or becomes extinct in both populations.

Critical migration rates. In general, the regions of stable infection polymorphism for the model with two-way migration could not be determined analytically. However, good approximations of these regions can be derived using the results of the previous section. In the mainland-island models, the cytotype that is dominant in the mainland population (i.e., either *Wolbachia* or "no infection") is favored in comparison to the model with two-way migration. Thus, if migration occurs in both directions, the spread of *Wolbachia* or "no infection" occurs at higher migration rates than in both mainland-island models; the minimum of the critical migration rates for the mainland-island models can therefore serve as a lower estimation for

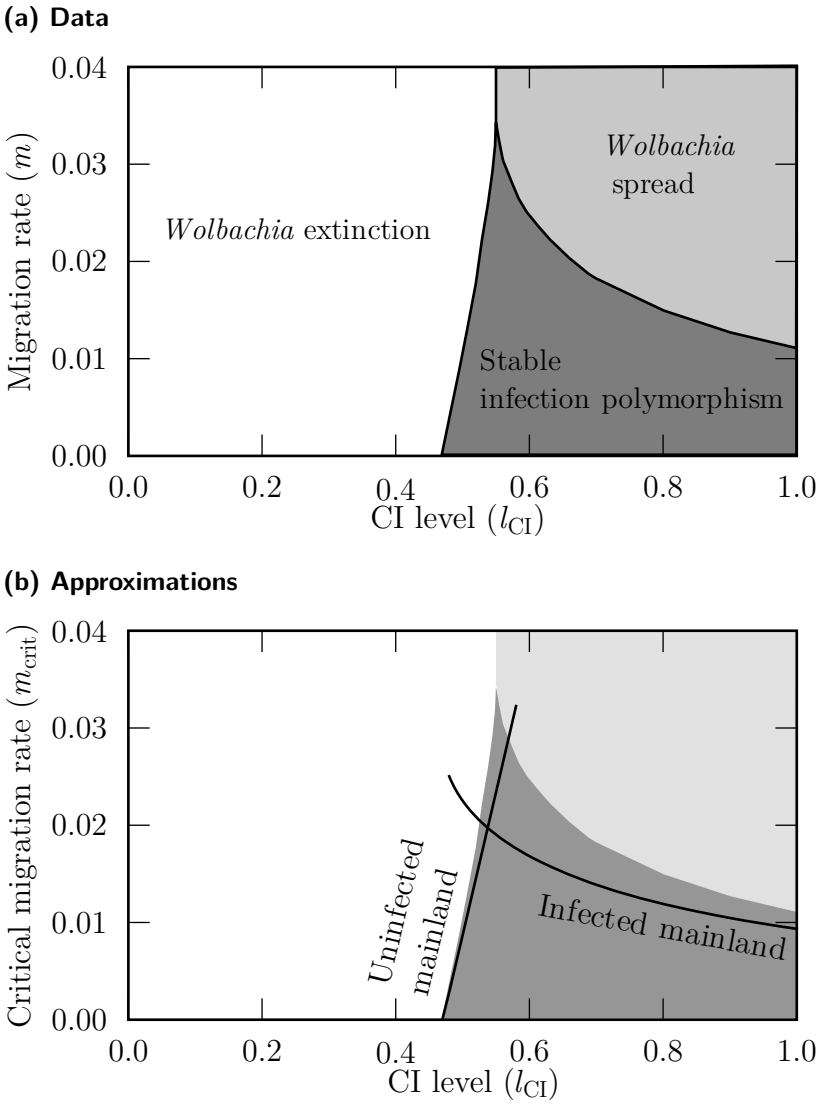


Figure 3.7: Stability of the infection polymorphism in the model with symmetric two-way migration. Subfigure (a) shows system equilibria determined by numerical simulations. There are three regions, (i) a region where the infection polymorphism is stable (shaded in dark gray), (ii) a region where *Wolbachia* spreads in both populations (light gray), and (iii) a region where *Wolbachia* goes to extinction in both populations (white). Subfigure (b) illustrates that the minimum of the analytical solutions of the critical migration rates derived for the two mainland-island models (black curves) can be used to approximate the region of stable infection polymorphism in the case of symmetric migration (compare Eq. (3.18)). Parameters: $f = 0.1$, $m = m_{\mathrm{A}} = m_{\mathrm{B}}$, $t = 0.9$.

the boundary of the region of stable infection polymorphism.

Symmetric migration. Using Equations (3.10) and (3.15), it holds that for symmetric migration, $m = m_A = m_B$, with given fecundity reduction, transmission rate, and level of CI (where $l_{\mathrm{CI}} > l_{\mathrm{crit}}$), a stable infection polymorphism is possible if $m \leq \min\{m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}},\ m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}\}$, i.e.,

$$m \leq \min\left\{ \frac{R(f, l_{\mathrm{CI}}, t)}{2\,(1-f)\,t\,D(f, l_{\mathrm{CI}}, t)}, \right. \\ \left. 2\,D(f, l_{\mathrm{CI}}, t)\left(\frac{1-\sqrt{(1-f)t}}{f + l_{\mathrm{CI}} + \sqrt{R(f, l_{\mathrm{CI}}, t)}}\right)^2 \right\}, \tag{3.18}$$

where $R(f, l_{\mathrm{CI}}, t)$ and $D(f, l_{\mathrm{CI}}, t)$ are the functions defined in Equations (3.3). This is illustrated in Fig. 3.7b. Note that the approximations become better with increasing transmission rates (compare e.g. Fig. 3.9a for the case of $t = 1$).

In general, the effect of increasing or decreasing parameter values on the infection polymorphism stability will depend on where in parameter space the system resides. E.g. if exceeding the critical migration rate results in a *Wolbachia* spread (the system moves from region (i) to region (ii) in terms of Inequalities (3.17) then usually any parameter change that strengthens *Wolbachia* (i.e., a decrease of fecundity reduction, or an increase of CI level or transmission rate) will render the infection polymorphism more unstable, whereas any parameter change that weakens *Wolbachia* (an increase of fecundity reduction, or a decrease of CI level or transmission rate) will improve its stability. Yet, if a parameter shift that weakens *Wolbachia* moves the system into the region where migration above the critical rate drives *Wolbachia* to extinction (i.e., a passage from region (i) to region (iii)) then any further change will have the exact opposite effect as just described. For instance, in Fig. (3.7a), when cytoplasmic incompatibility is complete, $l_{\mathrm{CI}} = 1$, the system resides in the parameter space region (i)/(ii). Decreasing the CI level weakens *Wolbachia* and yields increasing critical migration rates which can be interpreted as a stabilization of the infection polymorphism. But once the system makes the passage into region (i)/(iii), a further decrease of the CI level results in decreasing critical migration rates and hence destabilizes the infection polymorphism.

The largest critical migration rates and thus the highest stability are achieved for the parameter set $f = \frac{5}{9} = 55.\bar{5}\%$, $l_{\mathrm{CI}} = 1$, and $t = 1$, so that $m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}} = m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}} = \frac{1}{9} = 11.\bar{1}\%$. However, with fecundity reductions this severe, infection of any host population in the first place will be improbable because the threshold frequency for invasion of a host population is very high ($x_2^* = f = \frac{5}{9}$). For any given set of f and t values, the highest stability is obtained for intermediate CI levels that are close to the one for which $m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}} = m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}$ (compare Fig. (3.7b).

Asymmetric migration. The approach can be generalized for the case of asymmetric migration: If population A is uninfected and population B infected in the beginning, and if $l_{\mathrm{CI}} > l_{\mathrm{crit}}$, then the infection polymorphism is stable if the follow-

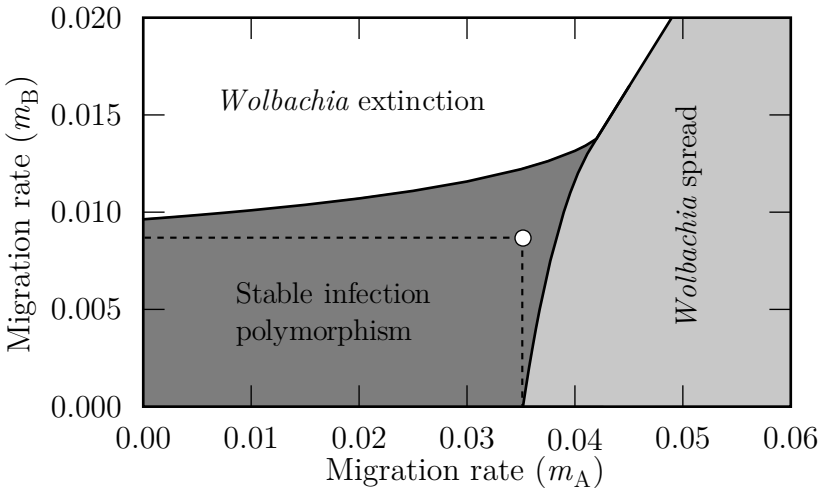


Figure 3.8: Stability of the infection polymorphism in the model with asymmetric two-way migration. System equilibria were determined by numerical simulations. As for the symmetric case, there are three regions, (i) a region where the infection polymorphism is stable (shaded in dark gray), (ii) a region where *Wolbachia* spreads in both populations (light gray), and (iii) a region where *Wolbachia* goes to extinction in both populations (white). The infection polymorphism is stable if migration rates stay within a certain region. This region can be approximated by the dashed rectangle. The upper right corner (open circle) is defined by the analytical solutions of the critical migration rates for the two mainland-island models (see Inequalities (3.19)). Parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.5$, $t = 0.9$.

ing two conditions hold:

$$m_A \leq m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}} = 2\,D(f, l_{\mathrm{CI}}, t)\left(\frac{1-\sqrt{(1-f)t}}{f + l_{\mathrm{CI}} + \sqrt{R(f, l_{\mathrm{CI}}, t)}}\right)^2, \tag{3.19a}$$

$$m_B \leq m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}} = \frac{R(f, l_{\mathrm{CI}}, t)}{2\,(1-f)\,t\,D(f, l_{\mathrm{CI}}, t)}. \tag{3.19b}$$

Again, $R(f, l_{\mathrm{CI}}, t)$ and $D(f, l_{\mathrm{CI}}, t)$ are the functions defined in Equations (3.3).

E.g., if we choose the parameters $f = 0.1$, $l_{\mathrm{CI}} = 0.5$, and $t = 0.9$, then the critical migration rates for the mainland-island models are $m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}} \approx 3.52\%$ if the mainland is uninfected and $m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}} \approx 0.87\%$ if the mainland is infected. Thus, from Inequalities (3.19) it follows that the infection polymorphism is stable for any pair (m_A, m_B) with $m_A \leq 0.87\%$ and $m_B \leq 3.52\%$. In Fig. 3.8, this region is illustrated by dashed lines, and the open circle marks the pair of migration rates $(0.87\%, 3.52\%)$. As can be seen in the graph, the whole range of stable infection polymorphism is slightly larger but can be approximated reasonably well in the described way.

We can use the area of the rectangle in the m_A–m_B plane that is defined by Inequalities (3.19) as an approximate measure for the stability of the infection polymorphism:

$$A(f, l_{\mathrm{CI}}, t) = m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}(f, l_{\mathrm{CI}}, t) \cdot m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}(f, l_{\mathrm{CI}}, t) \tag{3.20}$$

The larger the area, the more migration the system can endure while still maintaining the infection difference. The area of the rectangle is largest for the parameter set $f = \frac{3}{4}$, $l_{\mathrm{CI}} = 1$, and $t = 1$, with $m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}} = 6.25\%$ and $m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}} = 25\%$. However, as already argued in the case of symmetric migration, with fecundity reductions this severe, invasion of *Wolbachia* into any host population in the first place is improbable for this set of parameters due to a high invasion threshold. For given fecundity reduction and transmission rate, the system's infection polymorphism is most stable for complete CI (unlike the symmetric migration scenario, where intermediate levels yield the highest stability). This is because $m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}$ grows faster with increasing CI levels than $m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}$ decreases, which results in an ever enlarging product $A(f, l_{\mathrm{CI}}, t)$.

3.2.5 Special cases

We continue our analysis of infection polymorphism stability by considering three special cases of the full model. In each case, one of the parameters for fecundity reduction, strength of cytoplasmic incompatibility, and transmission rate will obey a boundary condition of its full range of values. In the first case, we assume that *Wolbachia* transmission is perfect, i.e., $t = 1$. Secondly, we set $f = 0$ and thus analyze the model without fecundity reductions. The analysis of these two cases has been published (Flor et al., 2007). Finally, the case of complete incompatibility, $l_{\mathrm{CI}} = 1$, will be examined. In each of these cases, we will merely present the simplified versions of dynamics, fixpoints, and critical migration rates for the single host population and the two mainland-island scenarios. For the two-way migration scenario, we will show corresponding results from numerical simulations. Each special case section is supposed to be comprehensible on its own without having read the full model Sections 3.2.1 to 3.2.4. If more detail is desired, however, all special case equations also refer to the equivalent full model equations.

3.2.5.1 Perfect transmission

In this section, we assume that *Wolbachia* transmission is perfect, $t = 1$, i.e., an infected female's eggs will all inherit the bacteria without exception. Offspring production follows the pattern shown in Tab. 3.3.

Single host population. If *Wolbachia* transmission is perfect, the dynamics in a panmictic host population (compare Eq. (3.1)) read

$$x' = F(f, l_{\mathrm{CI}}, 1, x) = \frac{(1-f)\,x}{1 - f\,x - l_{\mathrm{CI}}\,x\,(1-x)}. \tag{3.21}$$

The fixpoints (compare Equations (3.2)) greatly simplify to

$$x_1^* = 0, \tag{3.22a}$$

$$x_2^* = \frac{f}{l_{\mathrm{CI}}}, \tag{3.22b}$$

$$x_3^* = 1. \tag{3.22c}$$

If $f > 0$ then x_1^* is stable for any level of CI. Furthermore, if $l_{\mathrm{CI}} \leq l_{\mathrm{crit}}(f, 1) = f$ then $x_2^* \geq 1$, so that x_1^* is the only stable fixpoint in the frequency range $[0, 1]$,

and *Wolbachia* inevitably goes to extinction. But if $l_{\mathrm{CI}} > f$ then x_3^* is stable, and *Wolbachia* spreads to fixation from any starting frequency above the threshold frequency x_2^*.

Uninfected mainland. For the scenario where an island population receives migration from a mainland void of *Wolbachia*, the dynamics (compare Eq. (3.5)) and respective fixpoints (compare Equations (3.6)) are

$$x' = (1-m)\,F(f, l_{\mathrm{CI}}, 1, x) = \frac{(1-f)\,(1-m)\,x}{1 - f\,x - l_{\mathrm{CI}}\,x\,(1-x)} \tag{3.23}$$

and

$$x_1^* = 0, \tag{3.24a}$$

$$x_{2,3}^* = \frac{f + l_{\mathrm{CI}} \mp \sqrt{(l_{\mathrm{CI}} - f)^2 - 4\,(1-f)\,l_{\mathrm{CI}}\,m}}{2\,l_{\mathrm{CI}}}. \tag{3.24b}$$

The fixpoints x_2^* and x_3^* are real-valued and in the range $[0, 1]$ if

$$l_{\mathrm{CI}} > \hat{l}_{\mathrm{crit}}(f, m, 1) = f + 2\,(1-f)\,m + 2\sqrt{(1-f)\,m\,[f + (1-f)\,m]} \tag{3.25}$$

(compare Inequality (3.8a)). If this condition is satisfied, and the *Wolbachia* frequency is at x_3^*, then the infection polymorphism between mainland and island is stable.

The stability of the infection difference between an infected island and an uninfected mainland can also be described by a critical migration rate. If migration is larger than this critical value, the polymorphism breaks down because *Wolbachia*

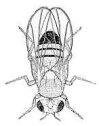
		♂		
				offspring infection status
♀		1	$1 - l_{\mathrm{CI}}$	uninfected
		$1 - f$	$1 - f$	infected

Table 3.3: Mating table for the model with perfect *Wolbachia* transmission. The relative number of offspring produced from different mating pairs depends on fecundity reductions, f, imposed on infected females and on the CI level, l_{CI}. E.g. if an uninfected female (♀) mates with a *Wolbachia (W)* infected male (♂) then a relative number $1 - l_{\mathrm{CI}}$ of uninfected offspring are produced.

goes to extinction. In the case of perfect transmission, the critical migration rate (compare Eq. (3.10)) is

$$m_{\text{crit}}^{\text{u}\to\text{w}}(f, l_{\text{CI}}, 1) = \frac{(l_{\text{CI}} - f)^2}{4\,(1-f)\,l_{\text{CI}}} \tag{3.26}$$

if $l_{\text{CI}} > f$; if $l_{\text{CI}} \leq f$, the critical migration rate is zero. In Fig. 3.9a, the critical migration rate is plotted as a function of the CI level for fixed fecundity costs, $f = 0.1$.

Infected mainland. When *Wolbachia* transmission from female to egg is perfect, *Wolbachia* in an infected host population is at fixation ($x_3^* = 1$; see Eq. (3.22c)). If migration into the island population occurs from a likewise infected mainland population, the dynamics (compare Eq. (3.12)) can be formulated as

$$x' = (1-m)\,F(f, l_{\text{CI}}, 1, x) + m = \frac{(1-f)\,x + m\,(1-x)\,(1-l_{\text{CI}}\,x)}{1 - f\,x - l_{\text{CI}}\,x\,(1-x)}, \tag{3.27}$$

provided that $l_{\text{CI}} > f$. The fixpoints of this system are (compare Equations (3.14))

$$x_1^* = 1, \tag{3.28a}$$

$$x_{2,3}^* = \frac{f + l_{\text{CI}}\,m \pm \sqrt{(f + l_{\text{CI}}\,m)^2 - 4\,l_{\text{CI}}\,m}}{2\,l_{\text{CI}}}. \tag{3.28b}$$

The stability of the infection polymorphism can again be assessed by a critical migration rate (compare Eq. (3.15)):

$$m_{\text{crit}}^{\text{w}\to\text{u}}(f, l_{\text{CI}}, 1) = \frac{\left(1 - \sqrt{1-f}\right)^2}{l_{\text{CI}}} \tag{3.29}$$

if $l_{\text{CI}} > f$. If $l_{\text{CI}} \leq f$ then the critical migration rate is zero. Figure (3.9a) shows how the critical migration rate decreases with increasing levels of CI (fecundity costs are fixed at $f = 0.1$).

Two-way migration. If migration between two host populations A and B is allowed to occur in both directions (compare Fig. 3.1 and Equations (3.16)), the system is not analytically tractable despite the simplification that results from perfect *Wolbachia* transmission. Hence, we conducted numerical simulations to screen the parameter space and find conditions which allow for a stable infection polymorphism. For a detailed description of the way the simulations were performed, see Section 3.2.4.

Fig. 3.9a shows the parameter space spanned by CI level and migration rate for $f = 0.1$ and $t = 1$ for symmetric migration, $m = m_A = m_B$, between two parapatric host populations. It also illustrates that the critical migration rates in the symmetric two-way migration model can be approximated reasonably well with the analytical solutions of the mainland-island critical migration rates (also

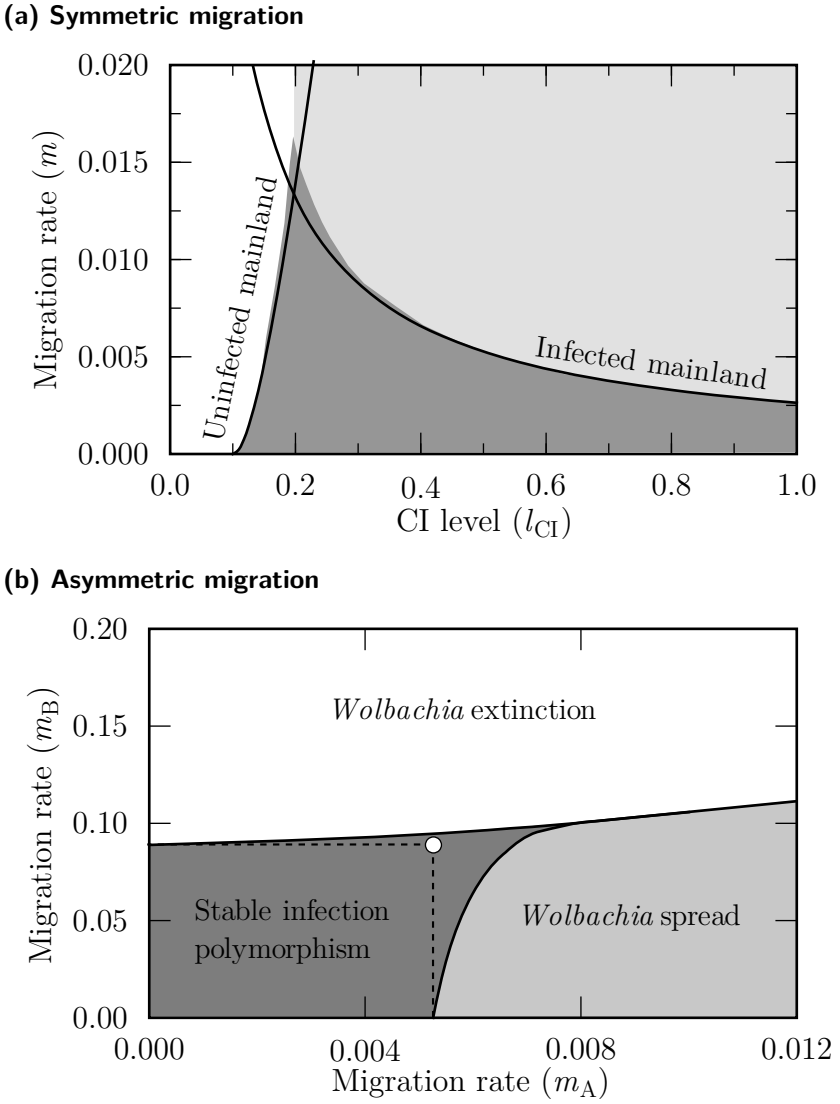


Figure 3.9: Stability of the infection polymorphism in the two-way migration model with perfect transmission. (a) Symmetric migration, $m = m_{\mathrm{A}} = m_{\mathrm{B}}$, and (b) asymmetric migration, $m_{\mathrm{A}} \neq m_{\mathrm{B}}$. Both subfigures show system equilibria determined by numerical simulations and analytical approximations. There are three regions, (i) a region where the infection polymorphism is stable (shaded in dark gray), (ii) a region where *Wolbachia* spreads to fixation in both populations (light gray), and (iii) a region where *Wolbachia* goes to extinction in both populations (white). In both subfigures, the analytical solutions of the critical migration rates derived for the two mainland-island models (black curves in (a) and dashed rectangle with open circle in the upper right corner in (b)) can be used to approximate the region of stable infection polymorphism (see Inequalities (3.30) and (3.31)). Parameters: (a) $f = 0.1$; (b) $f = 0.1$, $l_{\mathrm{CI}} = 0.5$

compare Inequality (3.18)):

$$m \leq \min\left\{ \frac{(l_{\mathrm{CI}} - f)^2}{4\,(1-f)\,l_{\mathrm{CI}}},\quad \frac{\left(1-\sqrt{1-f}\right)^2}{l_{\mathrm{CI}}} \right\}. \tag{3.30}$$

A similar approach can be used to approximate the region where the infection polymorphism is stable if migration rates are allowed to take different values in the two directions ($m_A \neq m_B$), i.e., for asymmetric migration (compare Inequalities (3.19):

$$m_A \leq m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}(f, l_{\mathrm{CI}}, 1) = \frac{\left(1-\sqrt{1-f}\right)^2}{l_{\mathrm{CI}}}, \tag{3.31a}$$

$$m_B \leq m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}(f, l_{\mathrm{CI}}, 1) = \frac{(l_{\mathrm{CI}} - f)^2}{4\,(1-f)\,l_{\mathrm{CI}}}. \tag{3.31b}$$

This is depicted in Fig. 3.9b. The dashed line rectangle corresponds to the region that is defined by Inequalities (3.31). The open circle demarks the point $(m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}},\ m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}) \approx (0.53\%,\ 8.\bar{8}\%)$. Note that the transmission rate increase from $t = 0.9$ to $t = 1$ in comparison to Fig. (3.8) results in an upward-leftward movement of the rectangle's upper right corner in the m_A–m_B plane. This will be true for any parameter change that yields a strengthening of the *Wolbachia* infection, just as a weakening of the infection leads to a downward-rightward movement.

3.2.5.2 No fecundity costs

In this section, we assume that the *Wolbachia* infection does not affect female fecundity ($f = 0$). Again, we briefly present the dynamics in a single host population, then discuss the mainland-island models and finally the model with two-way migration (compare Fig. 3.1). The production of offspring from the different mating pairs is shown in Tab. 3.4.

Single host population. If *Wolbachia* do not impose fecundity costs on female hosts, $f = 0$, then the infection dynamics within an isolated host population (compare Eq. (3.1)) become

$$x' = F(0, l_{\mathrm{CI}}, t, x) = \frac{t\,x}{1 - l_{\mathrm{CI}}\,x\,(1 - t\,x)}. \tag{3.32}$$

Because $R(0, l_{\mathrm{CI}}, t) = l_{\mathrm{CI}}^2 - 4\,l_{\mathrm{CI}}\,t\,(1-t)$ and $D(0, l_{\mathrm{CI}}, t) = 2\,l_{\mathrm{CI}}\,t$ (compare Equations (3.3)), the three fixpoints of Eq. (3.32) are:

$$x_1^* = 0, \tag{3.33a}$$

$$x_{2,3}^* = \frac{l_{\mathrm{CI}} \mp \sqrt{l_{\mathrm{CI}}^2 - 4\,l_{\mathrm{CI}}\,t\,(1-t)}}{2\,l_{\mathrm{CI}}\,t} \tag{3.33b}$$

(compare Equations (3.2)). If $t = 1$ then $x_1^* = x_2^* = 0$ is unstable, and $x_3^* = 1$ is stable. *Wolbachia* goes to fixation for any positive level of CI in that case. But if $t < 1$ then x_1^* is stable for any CI level and characterizes the state of *Wolbachia*

extinction. Furthermore, if $t < 1$ then *Wolbachia* can persist stably in the population if and only if $t > \frac{1}{2}$ and $l_{\mathrm{CI}} > l_{\mathrm{crit}}(0,t) = 4\,t\,(1-t)$. Then, x_2^* and x_3^* are real numbers, and it holds that $0 < x_2^* < x_3^* \leq 1$.

Uninfected mainland. We now take into account one-way migration from an uninfected mainland to an island population (see Figure 3.1a). Using function $F(0, l_{\mathrm{CI}}, t, x)$, we can formulate the infection dynamics in the island population as

$$x' = (1-m)\,F(0, l_{\mathrm{CI}}, t, x) = \frac{(1-m)\,t\,x}{1 - l_{\mathrm{CI}}\,x\,(1-t\,x)}. \tag{3.34}$$

Eq. (3.34) has the three fixpoints (compare Equations (3.6))

$$x_1^* = 0, \tag{3.35a}$$

$$x_{2,3}^* = \frac{l_{\mathrm{CI}} \mp \sqrt{l_{\mathrm{CI}}^2 - 4\,l_{\mathrm{CI}}\,t\,[1-(1-m)\,t]}}{2\,l_{\mathrm{CI}}\,t}. \tag{3.35b}$$

If $t < 1$ and $m > 0$ then x_1^* (the state of *Wolbachia* extinction) is stable for any level of CI. A stable persistence of *Wolbachia* on the island is possible if and only if $t > \frac{1}{2}$ and $l_{\mathrm{CI}} > \hat{l}_{\mathrm{crit}}(0, m, t) = 4\,t\,[1-(1-m)\,t]$.

We again apply the concept of the critical migration rate to analyze the stability of the infection polymorphism. The critical migration rate computes to

$$m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}(0, l_{\mathrm{CI}}, t) = \frac{l_{\mathrm{CI}} - 4\,t\,(1-t)}{4\,t^2} \tag{3.36}$$

if $l_{\mathrm{CI}} > 4\,t\,(1-t)$ and $t > \frac{1}{2}$. If $m < m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}$ and the infection frequency on

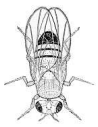
♀ \ ♂			offspring infection status
	1	$1 - l_{\mathrm{CI}}$	uninfected
	0	0	infected
	$1 - t$	$(1-t)\,(1-l_{\mathrm{CI}})$	uninfected
	t	t	infected

Table 3.4: Mating table for the model without fecundity costs. The relative number of offspring produced from different mating pairs depends on the CI level, l_{CI}, and the transmission rate, t. E.g. if *Wolbachia (W)* infected females (♀) and males (♂) mate then $(1-t)(1-l_{\mathrm{CI}})$ uninfected and t infected offspring are produced.

the island is at x_3^* then the infection difference between uninfected mainland and mainly infected island is maintained. If on the other hand migration occurs at a rate larger than $m_{\text{crit}}^{\text{u}\to\text{w}}$ then x_1^* is the only equilibrium frequency, and *Wolbachia* goes to extinction independent of its starting frequency, destroying the infection polymorphism as a result. In Fig. 3.10a, the critical migration rate is plotted as a function of the CI level for a fixed transmission rate, $t = 0.9$.

Infected mainland. The influence of migration from a mainland population where the infection is at the frequency $x_{\text{main}}(0, l_{\text{CI}}, t) = \frac{l_{\text{CI}}+\sqrt{l_{\text{CI}}^2-4\,l_{\text{CI}}\,t(1-t)}}{2\,l_{\text{CI}}\,t}$ (see Eq. (3.33b) on the *Wolbachia* dynamics in an island population can be modeled as

$$x' = (1-m)\,F(0, l_{\text{CI}}, t, x) + m\,x_{\text{main}}(0, l_{\text{CI}}, t) \tag{3.37}$$

(compare Eq. (3.12). These dynamics imply that $l_{\text{CI}} > 4\,t\,(1-t)$ and $t > \frac{1}{2}$. The three fixpoints of Eq. (3.37) are

$$x_1^* = x_{\text{main}}(0, l_{\text{CI}}, t) = \frac{l_{\text{CI}} + \sqrt{l_{\text{CI}}^2 - 4\,l_{\text{CI}}\,t\,(1-t)}}{2\,l_{\text{CI}}\,t}, \tag{3.38a}$$

$$x_{2,3}^* = \frac{1}{4\,l_{\text{CI}}\,t}\Bigg(l_{\text{CI}}(1+m) - (1-m)\sqrt{l_{\text{CI}}^2 - 4\,l_{\text{CI}}\,t\,(1-t)} \pm \sqrt{\left[l_{\text{CI}}(1+m) - (1-m)\sqrt{l_{\text{CI}}^2 - 4\,l_{\text{CI}}t(1-t)}\right]^2 - 16\,l_{\text{CI}}\,m\,t}\Bigg). \tag{3.38b}$$

If $m > 0$ and $t < 1$ then x_1^* is stable for all levels of CI, representing the spread of *Wolbachia* onto the island. A stable infection polymorphism is only possible if x_3^* is a real-valued stable fixpoint in the frequency range $[0, 1]$ This is the case if $l_{\text{CI}} > 4\,t\,(1-t)$, $t > \frac{1}{2}$, and if migration is below the critical migration rate, $m < m_{\text{crit}}^{\text{w}\to\text{u}}(0, l_{\text{CI}}, t)$. The critical migration rate is zero for $l_{\text{CI}} \leq 4\,t\,(1-t)$ or $t \leq \frac{1}{2}$. However, it is positive otherwise and calculates to

$$m_{\text{crit}}^{\text{w}\to\text{u}}(0, l_{\text{CI}}, t) = 4\,l_{\text{CI}}\,t\left(\frac{1-\sqrt{t}}{l_{\text{CI}} + \sqrt{l_{\text{CI}}^2 - 4\,l_{\text{CI}}\,t\,(1-t)}}\right)^2. \tag{3.39}$$

Fig. 3.10a shows how the critical migration rate depends on the level of CI level, if the transmission rate is fixed at $t = 0.9$.

Two-way migration. As described for the full model, migration between two parapatric populations in both directions can be incorporated by expanding the dynamical system to a set of two coupled difference Equations (see Equations (3.16) in Section 3.2.4) which is analytically not solvable. Numerical simulations corroborate that if $l_{\text{CI}} \leq 4\,t\,(1-t)$ or $t \leq \frac{1}{2}$, the *Wolbachia* infection gets lost in both populations even without migration. However, for $l_{\text{CI}} > 4\,t\,(1-t)$ and $t > \frac{1}{2}$, stable coexistence of an infected and an uninfected population is possible if migration rates stay within certain regions.

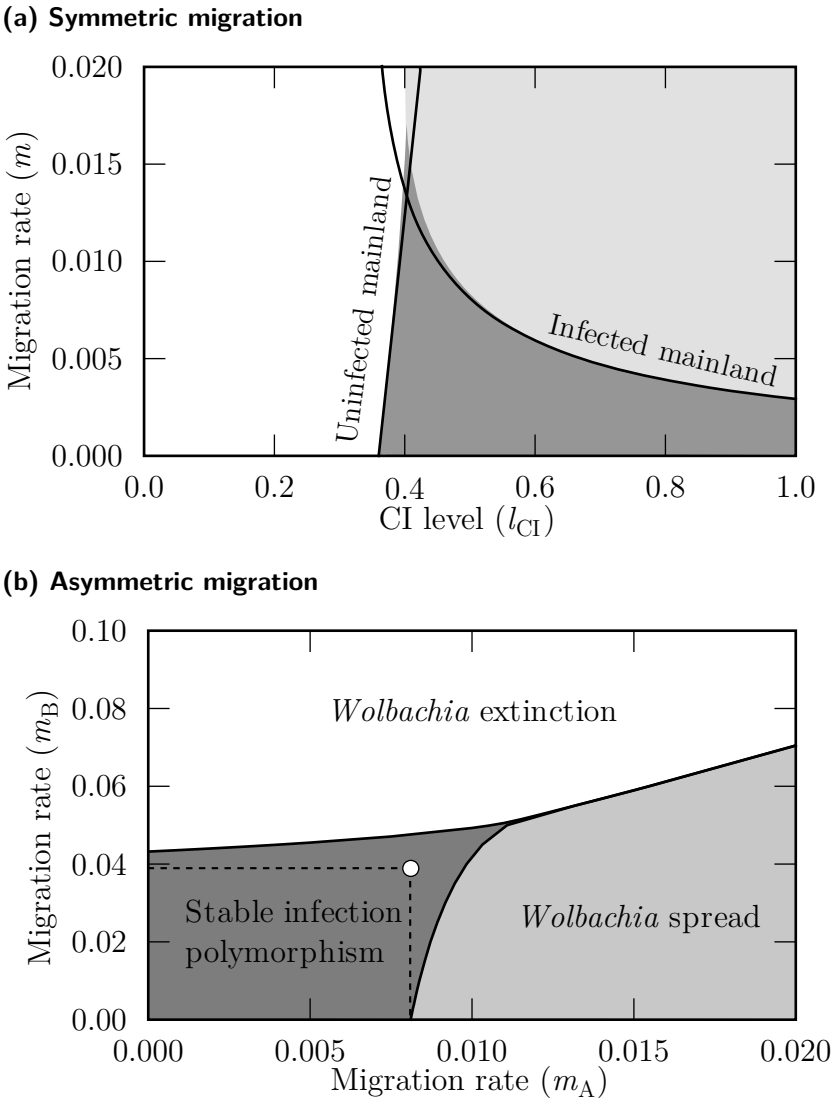


Figure 3.10: Stability of the infection polymorphism in the two-way migration model without fecundity costs. (a) Symmetric migration, $m = m_{\mathrm{A}} = m_{\mathrm{B}}$, and (b) asymmetric migration, $m_{\mathrm{A}} \neq m_{\mathrm{B}}$. System equilibria in both subfigures were determined by numerical simulations. There are three regions, (i) a region of stable infection polymorphism (shaded in dark gray), (ii) a region where *Wolbachia* spreads to the same frequency in both populations (light gray), and (iii) a region where *Wolbachia* becomes extinct in both populations (white). The analytical solutions of the critical migration rates derived for the two mainland-island models (solid black curves in (a) and dashed line rectangle with open circle in the upper right corner in (b)) can be used to approximate the parameter region of stable infection polymorphism (see Inequalities (3.40) and (3.41)). Parameters: (a) $t = 0.9$; (b) $l_{\mathrm{CI}} = 0.5$, $t = 0.9$

For symmetric migration rates ($m = m_A = m_B$), the parameter plane spanned by migration rate and CI level is shown in Fig. 3.10a (for transmission rate fixed at $t = 0.9$). *Wolbachia* goes to extinction if $l_{\mathrm{CI}} \leq 0.36$. But at larger CI levels the infection polymorphism is stable if the migration rate stays within the region

shaded in dark gray. Outside this region, *Wolbachia* goes to extinction in both populations for a rather narrow range of CI levels ($0.36 < l_{\mathrm{CI}} < 0.4$), and spreads from the infected to the uninfected population for high CI levels ($l_{\mathrm{CI}} > 0.4$).

The region of stable infection polymorphism can be approximated by using the analytical solutions of the critical migration rates for the mainland-island models, i.e., Eq. (3.36) and Eq. (3.39). Thus, for symmetric migration, it holds that an uninfected and an infected host population can stably coexist if

$$m \leq \min \left\{ \frac{l_{\mathrm{CI}} - 4t\,(1-t)}{4t^2},\ \frac{4t\left(1+t-2\sqrt{t}\right)}{\left[\sqrt{l_{\mathrm{CI}}} + \sqrt{l_{\mathrm{CI}} - 4t\,(1-t)}\right]^2} \right\}, \tag{3.40}$$

provided that $l_{\mathrm{CI}} > 4\,t\,(1-t)$ and $t > \frac{1}{2}$ (see black curves in Fig. 3.10a).

Similarly, the conditions under which a stable infection polymorphism is possible in the general case of asymmetric migration, can be approximately described as:

$$m_A \leq \frac{4t\left(1+t-2\sqrt{t}\right)}{\left[\sqrt{l_{\mathrm{CI}}} + \sqrt{l_{\mathrm{CI}} - 4t\,(1-t)}\right]^2}, \quad \text{and} \tag{3.41a}$$

$$m_B \leq \frac{l_{\mathrm{CI}} - 4t\,(1-t)}{4t^2}, \tag{3.41b}$$

where again $l_{\mathrm{CI}} > 4\,t\,(1-t)$ and $t > \frac{1}{2}$, and where population A is uninfected and population B infected in the beginning (compare Inequalities (3.19)). In Fig. 3.10b, these conditions are depicted by the dashed line rectangle.

3.2.5.3 Complete incompatibility

As the last special case, we summarize the dynamical behaviour of our model under the assumption that the cytoplasmic incompatibility is complete, i.e., sperm modification by *Wolbachia* is so efficient as to completely prohibit the fertilization of uninfected eggs. In our mathematical model, this is expressed by letting $l_{\mathrm{CI}} = 1$. Tab. 3.5 demonstrates how this affects the production of offspring through the different mating pairs.

Single host population. For complete CI levels, $l_{\mathrm{CI}} = 1$, the infection dynamics in a panmictic host population (compare Eq. (3.1)) are:

$$x' = F(f,1,t,x) = \frac{(1-f)\,t\,x}{1 - f\,x\,[1-(1-t)\,x] - x\,(1-t\,x)}. \tag{3.42}$$

The functions R and D (compare Equations (3.3)) simplify to $R(f,1,t) = (1-f)^2\,(1-2\,t)^2$ and $D(f,1,t) = 2\,[f+(1-f)\,t]$. So do the fixpoints (compare Equations (3.2)):

$$x_1^* = 0, \tag{3.43a}$$

$$x_2^* = \frac{1-(1-f)t}{f+(1-f)t}, \tag{3.43b}$$

$$x_3^* = 1. \tag{3.43c}$$

The fixpoint x_1^* is stable for all valid parameter conditions (compare Tab. 3.1). For the second fixpoint, it holds that $0 < x_2^* < 1$ if $t > \frac{1}{2}$. If this condition is satisfied then x_3^* is stable, and the infection can stably persist in the population. Otherwise, the system converges to x_1^*, and *Wolbachia* goes to extinction for any starting frequency, $x < 1$. Note that even though transmission rates are imperfect, and hence uninfected eggs will always be produced, *Wolbachia* spreads to fixation if the invasion threshold, x_2^*, is overcome. This is due to the fact that if all males are infected, *Wolbachia*-free eggs can not be fertilized at all, and no uninfected offspring is produced (also see Tab. 3.5).

Uninfected mainland. With migration of solely uninfected hosts from a mainland population, the infection dynamics in an island population (compare Eq. (3.5)) can be formulated as

$$x' = (1-m)\,F(f,1,t,x) = \frac{(1-f)\,(1-m)\,t\,x}{1-f\,x\,[1-(1-t)\,x]-x\,(1-t\,x)}. \tag{3.44}$$

The fixpoints in this case are (compare Equations (3.6)):

$$x_1^* = 0, \tag{3.45a}$$

$$x_{2,3}^* = \frac{1+f \mp \sqrt{(1-f)^2 - 4\,(1-f)t\big[m+(1-f)(1-m)(1-t)\big]}}{2\,\big[f+(1-f)t\big]}. \tag{3.45b}$$

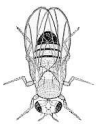
♀ \ ♂			offspring	infection	status
	1	0			uninfected
	0	0			infected
	$(1-f)(1-t)$	0			uninfected
	$(1-f)\,t$	$(1-f)\,t$			infected

Table 3.5: Mating table for the model with complete incompatibility. The relative number of offspring produced from different mating pairs depends on the fecundity costs imposed on infected females, f, and the transmission rate, t. Sperm of *Wolbachia (W)* infected males (♂) is only compatible with eggs of infected females (♀) that have inherited the infection from their mothers.

The fixpoint x_1^* is stable for any parameter constellation and represents an island void of *Wolbachia*. The fixpoints $x_{2,3}^*$ are real numbers in the range $[0,1]$ if $t > \frac{1}{2}$, and if the migration rate is below the critical migration rate (compare Eq. (3.10)):

$$m_{\text{crit}}^{\text{u}\to\text{w}}(f,1,t) = \frac{(1-f)\,(1-2\,t)^2}{4\,t\,[f+(1-f)\,t]}. \tag{3.46}$$

Infected mainland. With CI being complete, *Wolbachia* is at fixation in an infected population. Hence, the infection dynamics in an island population that receives migration from an infected mainland (compare Eq. (3.12)) are as follows:

$$\begin{aligned} x' &= (1-m)\,F(f,1,t,x) + m \\ &= \frac{(1-f)\,t\,x + m\,(1-x)\,(1-x\,[f+(1-f)\,t])}{1-x\,(1+f-x\,[f+(1-f)\,t])}. \end{aligned} \tag{3.47}$$

Letting $x' = x \equiv x^*$ yields the fixpoints (compare Equations (3.14)) of Eq. (3.47):

$$x_1^* = 1, \tag{3.48a}$$

$$\begin{aligned} x_{2,3}^* = {} & \frac{f+(1-f)\,t+m\,[1-(1-f)\,t]}{2\,[f+(1-f)\,t]} \\ & \pm \frac{\sqrt{(f+(1-f)\,t+m\,[1-(1-f)\,t])^2 - 4\,m\,[f+(1-f)\,t]}}{2\,[f+(1-f)\,t]}. \end{aligned} \tag{3.48b}$$

The critical migration rate reduces enormously in the case of complete CI levels (compare Eq. (3.15)):

$$m_{\text{crit}}^{\text{w}\to\text{u}}(f,1,t) = \frac{\left(1-\sqrt{(1-f)\,t}\right)^2}{f+(1-f)\,t}. \tag{3.49}$$

Two-way migration. To complete our analysis of the special cases, we end the results section by stating the conditions that approximately describe the infection polymorphism stability when migration between two populations occurs in both directions in the case of complete cytoplasmic incompatibility. For symmetric migration, the infection difference may be stable if the migration rate does not exceed a critical threshold (compare Inequality (3.18)):

$$m \le \min\left\{ \frac{(1-f)\,(1-2\,t)^2}{4\,t\,[f+(1-f)\,t]},\ \frac{\left(1-\sqrt{(1-f)\,t}\right)^2}{f+(1-f)\,t} \right\}. \tag{3.50}$$

If asymmetric migration is considered, a (mainly) uninfected and a (mainly) infected host population can persist in parapatry if the migration rates satisfy the following conditions (compare Inequalities (3.19)):

$$m_A \le \frac{\left(1-\sqrt{(1-f)\,t}\right)^2}{f+(1-f)\,t}, \quad \text{and} \tag{3.51a}$$

$$m_B \leq \frac{(1-f)\,(1-2\,t)^2}{4\,t\,[f+(1-f)\,t]}. \tag{3.51b}$$

3.2.6 Local adaptation

To conclude our analysis, we will briefly discuss how local selection at a trait locus, T, of the *Wolbachia* host influences the stability of an infection polymorphism. In order to do this, we consider an island population that receives migration from a mainland population that is either infected with *Wolbachia* or uninfected. In the island population, a trait allele T_1 is favored by selective forces before other traits at the T locus. In particular, hosts carrying allele T_1 have an increased viability by a factor of $1+s$ compared to hosts harboring allele T_2 which is fixed in the mainland population. The selection coefficient, s, takes non-negative values, $s \geq 0$. For a full description of the mathematical model, see Section 5.2. The model can not be solved analytically. However, our numerical simulations show that in general, the critical migration rate increases with stronger selection at the host trait locus. This is because local adaptation favors residents in comparison to migrants, and the infection type and selected locus tend to be coupled in association disequilibrium, which imparts a selective advantage to both in the resident population. Fig. 3.11 shows critical migration rates (a) for both mainland-island scenarios and (b) for symmetric two-way migration. Local adaptation can result in comparatively high critical migration rates. If, for example, the CI level is $l_{\mathrm{CI}} = 0.9$, fecundity costs are $f = 0.1$, and *Wolbachia* is transmitted at a rate $t = 0.9$, then the critical migration rates for $s = 0$ are $m_{\mathrm{crit}} \approx 1.3\%$ for the infected mainland scenario and $m_{\mathrm{crit}} \approx 14.2\%$ for the uninfected mainland scenario; but for $s = 1$ the rates are $m_{\mathrm{crit}} \approx 3.7\%$ and $m_{\mathrm{crit}} \approx 29.0\%$, respectively. Note that $s = 1$ corresponds to a two-fold fitness advantage for the resident trait allele. This shows that local adaptation significantly stabilizes postzygotic isolation induced by unidirectional CI.

It is possible to derive analytical lower estimations for the critical migration rates with local host adaptation. In order to do so, we make the following simplifying assumptions regarding the trait locus and nuclear inheritance thereat: Let all migrants carry trait T_2, and let all residents on the island display trait T_1. Furthermore, let all offspring inherit the locally adaptive trait, T_1. Note that because the latter assumption is beneficial to the migrants, the spread of the mainland cytotype on the island is facilitated.

Denoting the *Wolbachia* frequency in consecutive generations by x and x', the infection dynamics of this system can be written as

$$x^+ = \frac{(1-m)\,(1+s)\,x + m\,x_{\mathrm{main}}}{1+(1-m)\,s}, \tag{3.52a}$$

$$x' = F(f, l_{\mathrm{CI}}, t, x^+), \tag{3.52b}$$

where x^+ denotes the *Wolbachia* frequency after migration and viability selection, where x_{main} is the *Wolbachia* frequency in the mainland population, and where $F(f, l_{\mathrm{CI}}, t, x)$ describes the reproduction step as derived in Eq. (3.1). If the mainland is uninfected then $x_{\mathrm{main}} = 0$; and if the mainland is infected then $x_{\mathrm{main}} = x_{\mathrm{main}}(f, l_{\mathrm{CI}}, t)$, according to Eq. (3.11). For both cases, the fixpoints x_1^*,

x_2^*, and x_3^*, can be calculated analytically (see Equations (A.1) and (A.2) in the appendix section A). The same applies to the critical migration rate which is obtained by assuming $x_2^* = x_3^*$. In the two following equations, the functions $D(f, l_{\mathrm{CI}}, t)$ and $R(f, l_{\mathrm{CI}}, t)$ are given by Equations (3.3). For the case of an uninfected mainland population, the critical migration rate is

$$m_{\mathrm{crit}}^{\mathrm{u}\to\mathrm{w}}(f, l_{\mathrm{CI}}, s, t) = \frac{(1+s)\, R(f, l_{\mathrm{CI}}, t)}{2\,(1-f)\, t\, D(f, l_{\mathrm{CI}}, t) + s\, R(f, l_{\mathrm{CI}}, t)}. \tag{3.53}$$

Note that for $s = 0$, the critical migration rate derived in Section 3.2.2 is reproduced (see Eq. (3.10)). If the mainland population is infected with *Wolbachia* elaborate rearrangement of $x_2^* = x_3^*$ shows that the critical migration rate is

$$m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}(f, l_{\mathrm{CI}}, s, t) = \frac{(1+s)\left[2\, D(f, l_{\mathrm{CI}}, t)\left(1 - \sqrt{(1-f)t}\right)^2 + s\left(f + l_{\mathrm{CI}} - \sqrt{R(f, l_{\mathrm{CI}}, t)}\right)^2\right]}{\left((f + l_{\mathrm{CI}})(1-s) + (1+s)\sqrt{R(f, l_{\mathrm{CI}}, t)}\right)^2 + 8\, s\, D(f, l_{\mathrm{CI}}, t)}. \tag{3.54}$$

Again, the critical migration rate of the corresponding model without local adaptation (see Eq. (3.15)) is obtained by letting $s = 0$. In Fig. 3.11a, both critical migration rates are shown as functions of the selection coefficient. The other parameters are fixed at $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, and $t = 0.9$.

If migration is above the critical rate, the mainland cytotype will spread onto the island. Thus, if the mainland is uninfected, *Wolbachia* will go to extinction, whereas it will spread if the mainland is infected. The infection polymorphism can be maintained only if migration is below the critical migration rate. As argued above, the spread of the mainland cytotype on the island is facilitated in the simplified system as compared to the full system with explicit modeling of nuclear inheritance (see Section 5.2 in Chapter 5). Therefore, it holds that the critical migration rates in Equations (3.53) and (3.54) represent lower estimates of the critical migration rates for the full system. Furthermore, they can be used as lower approximations of the critical migration rates in a scenario with two-way migration, as described in Section 3.2.4.

It is interesting to note that for the scenario with an uninfected mainland, the analytical approximation of the critical migration rate overestimates the true critical migration rate if selection is weak (see Fig. 3.11a). The same effect shows for the other mainland-island scenario but only for very small selection coefficients so that it is not apparent from the plot. To understand this effect, one has to bear in mind that the island trait is subject to a selection-migration balance. Roughly put, the locally adaptive trait can only be maintained on the island if $s > m$. For the stability of the infection polymorphism, this means that the coefficient of local selection must be larger than the critical migration rate (without local adaptation, i.e., according to Equations (3.10) and (3.15)) for our approximations to be applicable. Otherwise, selection at the trait locus is so weak that the trait allele that is fixed in the mainland population spreads onto the island and supersedes the locally adaptive trait. In that case, the infection polymorphism can obviously not be affected any longer by local

selective forces, and our approximations produce an overestimate of the true critical migration rate.

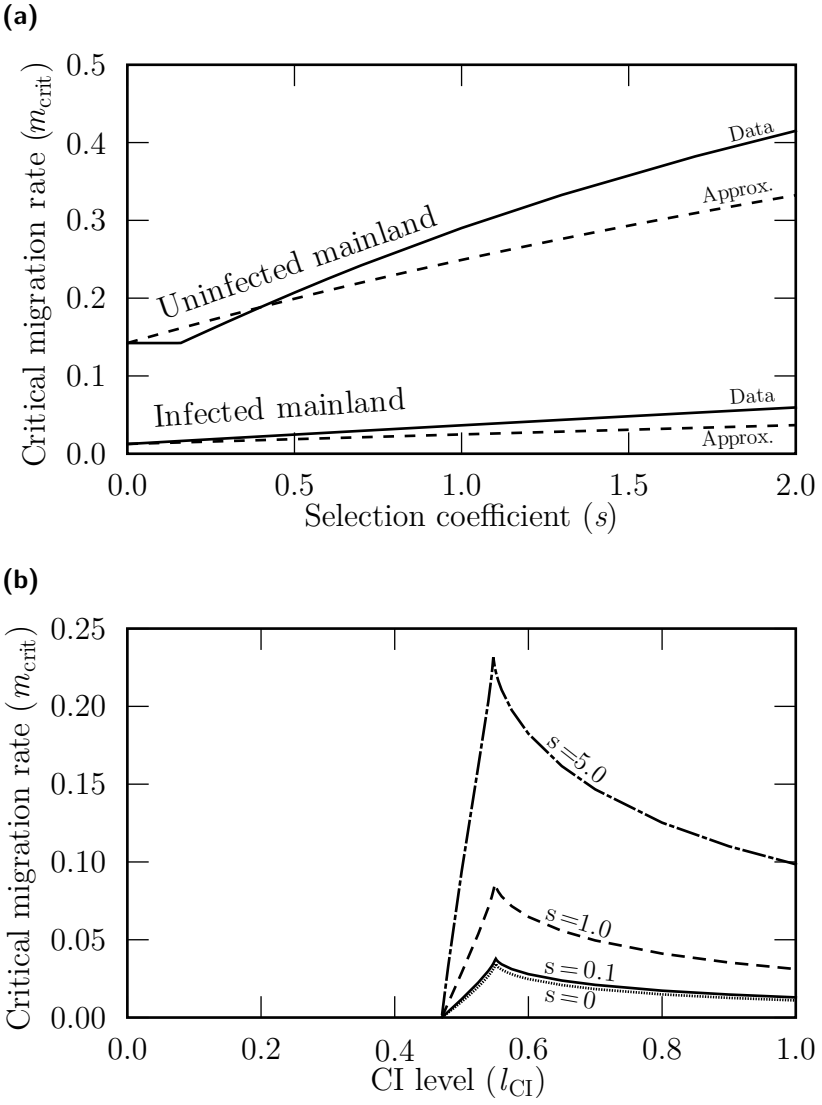


Figure 3.11: Critical migration rates for the model with local host adaptation. With local selection acting on a trait locus in the host, critical migration rates increase. This applies to all classes of population structure. In subfigure (a), the critical migration rate is shown as a function of the selection coefficient, s, for the two mainland-island scenarios. Solid lines represent data from numerical simulations, dashed lines are analytical approximations (compare Inequalities (3.53) and (3.54)). In subfigure (b), the symmetric two-way critical migration rate is plotted as a function of the CI level for different strengths of local selection (data from numerical simulations). Parameters: (a) $f = 0.1$, $l_{\text{CI}} = 0.9$, $t = 0.9$; (b) $f = 0.1$, $t = 0.9$.

3.3 Discussion

In this chapter, we investigated the dynamics of *Wolbachia*-induced unidirectional CI in a spatial model with two host populations linked by migration. Our main result is that (mainly) infected and (mainly) uninfected host populations can stably coexist if migration is below a critical migration rate and if the infection is costly or imperfectly transmitted. Under these circumstances, unidirectional CI acts as a postzygotic isolation mechanism between the populations. We determined the critical migration rates analytically for mainland-island models and showed that these solutions are lower estimations for the general model with two-way migration. Note that in the case of symmetric two-way migration, an infection polymorphism is most stable for intermediate levels of CI. By contrast, for asymmetric migration the infection difference is most stable for complete CI.

Previous studies on the dynamics of unidirectional CI have focused either on a single panmictic host population (Caspari and Watson, 1959; Fine, 1978; Turelli, 1994) or on the spatial spread of *Wolbachia* (Turelli and Hoffmann, 1991; Wade and Stevens, 1994; Schofield, 2002). Our results point out that population structure can prevent *Wolbachia* from spreading, and suggest that *Wolbachia* mosaics with infected and uninfected populations close by may exist in nature, especially if levels of CI are intermediate, or if migration is asymmetric e.g. due to predominant wind directions.

Such mosaically structured infection patterns might strongly influence the evolution of both *Wolbachia* and the host. This is due to unidirectional CI causing an asymmetric gene flow reduction between infected and uninfected populations (see Telschow et al., 2002a,b, 2007, and the next Chapter 4). Because of this gene flow distortion, infected populations are converted into population genetic sinks. Local adaptations are therefore favored in uninfected host populations compared to infected ones, while at the same time local host adaptations tend to enable more stable mosaic patterns. Further, host adaptation to *Wolbachia* is impeded. This might have a huge impact on the coevolution of host and symbiont. For male-killing inducing *Wolbachia*, it was shown theoretically that asymmetric gene flow can prevent adaptation in a host population infected by a male-killer (Telschow et al., 2006).

Our results support the view that unidirectional CI could be a factor in *Wolbachia* host speciation. This is because unidirectional CI causes postzygotic isolation in hybrid zones between infected an uninfected host populations. Postzygotic isolation, however, is widely believed to play a crucial role in speciation because it presumably selects for female mating preference (see Coyne and Orr, 2004, for a review). Thus, in uninfected host populations that receive migration from infected populations, unidirectional CI might select for premating isolation by reinforcement of female mating preferences if the infection polymorphism is stable, and thus facilitate speciation (see Chapters 5 and 6 and Shoemaker et al., 1999; Jaenike et al., 2006; Telschow et al., 2007).

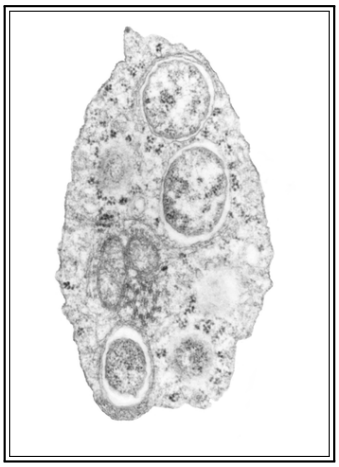
From a theoretical point of view, our model approach is a natural extension of single population models. Fine (1978) has shown that in order for CI to spread in a single host population, *Wolbachia* has to be at a frequency above a certain threshold. The critical migration rate introduced in this study is the two-population analogon of Fine's threshold frequency. *Wolbachia* spreads from one population to

another only if migration is above a threshold, the critical migration rate. Previous theoretical studies on spatial CI dynamics have not described such a threshold (Hoffmann and Turelli, 1997; Schofield, 2002). The reason is that these models are based on partial differential equations with migration characterized by a diffusion term. These approaches, however, do not incorporate enough population structure to observe the effects described in this study.

Symbiotic bacteria have been proposed as means to introduce useful genes in insect pest populations (Beard et al., 1993; Sinkins et al., 1997). For example in insect disease vectors, such useful genes may render the insect refractory to infection with a pathogenic agent, or reduce the competence to transmit these agents. With the self-spreading power of unidirectional CI, *Wolbachia* could provide the vehicle for the introduction of foreign genes in targeted arthropods. Our results imply that such measures can only then be effectively applied if migration between host populations generally occurs with rates above the critical migration rate. For biological pest control, high levels of CI are desirable because propagation of useful genes within host populations will be fastest under these conditions, but also because critical migration rates are low for CI levels close to one, facilitating the spread between populations. Thus, knowing the spatial structure of pest populations may prove essential for a successful application of this mechanism.

Previously, it was shown that *Wolbachia*-induced bidirectional CI is stable in face of two-way migration at high rates of up to 15% per generation (Telschow et al., 2005b). The critical migration rates observed for unidirectional CI are roughly one order of magnitude lower. However, even for unidirectional CI, high critical migration rates might be observed if additional genetical or ecological factors are included in the model. In the case of bidirectional CI, it was shown that local adaptations in the host can significantly increase the critical migration rates (Telschow et al., 2005a). This is because linkage disequilibria build up that stabilize both infection differences and local adaptations. We have demonstrated that the same stabilizing effect is effective in systems with unidirectional CI. Other reasonable scenarios resulting in higher critical migration rates are (i) local selection against *Wolbachia* due to naturally occurring antibiotics and (ii) local infections with a *Wolbachia* strain that causes a sex ratio distortion rather than CI (Engelstädter et al., 2004).

In summary, our results suggest that a stable coexistence between infected and uninfected parapatric host populations is possible under biologically reasonable conditions. We analyzed the role of migration in detail and showed that, if migration is below a critical value, then unidirectional CI acts as a postzygotic isolation mechanism. These results might have important implications for host evolution including speciation, for the coevolution of *Wolbachia* and its hosts, and for utilization of CI-inducing *Wolbachia* in biological pest control.



Chapter 4

Unidirectional cytoplasmic incompatibility and gene flow reduction

Stable patterns of parapatric *Wolbachia*-infected and uninfected host populations act as postzygotic reproductive isolation barriers between the populations if *Wolbachia* induces cytoplasmic incompatibility (CI). This effect can be attributed to the reduction of gene flow. We analyze the extent of gene flow modification in a mathematical model and present analytical approximations of the effective migration based on reproductive values of immigrants. Three different scenarios are investigated: Migration into a population with a *Wolbachia* infection at fixation, migration into a host population void of *Wolbachia*, and migration into a population with uninfected and infected hosts coexisting. We show that the gene flow is reduced asymmetrically between (mainly) infected and (mainly) uninfected populations and confirm the quality of our estimations by numerical simulations. The results are discussed in the context of local adaptations of the host and *Wolbachia*-driven speciation scenarios.

4.1 Introduction

The biological species concept (BSC) provides the most commonly used definition of species in evolutionary biology. It considers species to be groups of populations reproductively isolated from other such groups by "isolating mechanisms". According to the BSC, speciation is the evolution of barriers to gene exchange between such groups of populations of one species (Dobzhansky, 1935, 1937; Mayr, 1942). The amount of gene flow between populations hence indicates the degree of a speciation process; ultimately gene flow must be negligible between species.

Gene flow between populations can often be attributed to migration. However, if reproductive success of migrants in a new population is affected by various factors (such as genetic incompatibilities, divergent selection, etc.), *effective* migration may well differ from real migration. The effective migration rate is well-defined in various contexts and can therefore be used to determine and compare the intensities of gene flow between populations (e.g., Bengtsson, 1985; Barton and Bengtsson, 1986; Gavrilets, 1997; Rousset, 1999; Telschow et al., 2002b). To make the measure independent of the amount of real migration, Bengtsson (1985) used the ratio of effective migration rate, m_{eff}, to real migration rate, m, and coined the term *gene flow factor*: $gff = m_{\text{eff}}/m$. If $gff < 1$ (i.e., $m_{\text{eff}} < m$) then gene flow is reduced, whereas it is enhanced if $gff > 1$ ($m_{\text{eff}} > m$). Note that if one considers relative contributions of several reproductive isolation barriers to overall gene flow, it is possible to restate Eq. (2.1) simply as

$$G = \prod_{i=1}^{n} (1 - I_i) = \prod_{i=1}^{n} gff_i. \tag{4.1}$$

Here, I_i and gff_i are, respectively, the strength and gene flow factor of barrier i in a series of n barriers acting sequentially over the life cycle of the organisms under investigation.

In this study, we follow Kobayashi and Telschow (2008) and use a definition of the effective migration rate that was introduced to measure gene flow in and between class-structured populations. In the context of host populations with a stable *Wolbachia* polymorphism, one can formulate this definition as follows: Considering the frequency dynamics at a selectively neutral gene locus, the effective migration rate between two populations with a stable *Wolbachia* polymorphism (as in Chapter 3) is defined as the migration rate that - in a scenario without *Wolbachia* - would result in the same equilibrium frequencies at that locus (see also Section 4.2.5). We refer to that neutral locus as the "marker" locus. Note that a scenario where both populations are infected yields the same divergence at the marker locus as a scenario without *Wolbachia*, and the effective migration rate is equal to the real migration rate.

For bidirectional CI, i.e., for a scenario where two populations are infected with incompatible *Wolbachia* strains, Telschow et al. (2002a) derived an analytical approximation for the effective migration rate. Telschow et al. (2007) then applied their approach to the case of unidirectional CI, namely for *Wolbachia* infected immigrants in an uninfected population. Under the assumption that the migration rate is small the effective migration rate between an infected mainland and an uninfected

island is in good approximation

$$m_{\text{eff}} \approx m\frac{1-l_{\text{CI}}}{1+f}. \tag{4.2}$$

This formula was derived by considering the matriline of female migrants over successive generations and tracking the reproductive costs that arise from CI (Telschow et al., 2007). This may seem counterintuitive because only infected males suffer from CI, but the recursive nature of reduction in effective migration rate shows that gene flow is inhibited in the matriline too because each generation sons of female migrants inherit the infection and therefore suffer from CI, thus further reducing input of genes from immigrant descendants into the population. Telschow et al. (2002a) made the assumption that all residents of a host population share the same cytotype. In this chapter, we overcome this shortcoming and generalize the concept of effective migration between two populations with a stable *Wolbachia* polymorphism, along the way putting the approach mathematically on a more solid basis by building on a framework developed by Kobayashi et al. (2008) and Kobayashi and Telschow (2008).

In a limit theorem, Kobayashi et al. (2008) proved that there is a close connection between the effective migration rate, m_{eff}, and the average reproductive value of immigrants, $\bar{\nu}$:

$$\lim_{m\to 0}\frac{m_{\text{eff}}}{m} = \bar{\nu}. \tag{4.3}$$

Furthermore, they showed that under weak migration, i.e., if $m \ll 1$, the effective migration rate is in good approximation equal to the real migration rate weighed with the average reproductive value of immigrants:

$$m_{\text{eff}} \approx \bar{\nu}\ m. \tag{4.4}$$

Note that in the context of the unidirectional CI scenarios that this work is concerned with, weak migration is usually required in order for the *Wolbachia* infection polymorphism not to be destroyed (see Chapter 3).

The concept of reproductive value is a classical one in evolutionary biology and traces back to Fisher (1930). In a class-structured population, the reproductive value of a class can be defined as the normalized probability that a gene (on a neutral marker locus) drawn at random from the population in the distant future originates from an individual of the focal class (Fisher, 1930; Taylor, 1990; Grafen, 2006). The normalization ensures that the average reproductive value of the population is equal to one. In other words, the reproductive value is a measure for the relative contribution of a focal class to the future gene pool. The reproductive value of an individual host that belongs to a certain class is the per-class reproductive value divided by the class frequency. This individual reproductive value is also called *per-capita* reproductive value, and it is noteworthy that it is equal to the sum of the reproductive values of the individual's offspring.

The fitness graph method was recently introduced by Kobayashi and Telschow (2008) to allow for a relatively convenient way to calculate reproductive values in class-structured populations. As a closed graph, it appealingly presents dynamical relations between all classes of a population, and —once set up— it is straightforward to translate the fitness graph into a system of coupled equations or a fitness

matrix.

In this chapter, we show that a stable *Wolbachia* polymorphism asymmetrically reduces gene flow between host populations. We analytically calculate reproductive values of immigrants based on *Wolbachia's* dynamical parameters, and estimate the gene flow factor of the cytoplasmic reproductive barrier based on Eq. (4.4). We further analyze by numerical simulations the goodness of this approximation.

4.2 Model and results

As in the previous Chapter 3, we will make use of the three parameters that are commonly used to describe *Wolbachia* dynamics: f (fecundity costs incurred by infected females), l_{CI} (level of cytoplasmic incompatibility), and t (transmission fidelity of *Wolbachia*) (see Table 3.1 and Section 3.2 for details). Host generations are discrete and non-overlapping. Individuals reproduce sexually with a primary sex ratio of 1:1.

Based on the infection dynamics, we apply the fitness graph approach introduced by Kobayashi and Telschow (2008) in order to calculate reproductive values of uninfected and *Wolbachia* infected immigrants. We use the calculated reproductive values to estimate effective migration rates based on Eq. (4.4) and show by means of numerical simulations that this approach indeed yields good approximations of the effective migration rate and thus can be used to gauge gene flow between host populations with divergent infection states.

Previous authors have used the fitness matrix and fitness tree approaches to calculate effective migration rates (e.g., Bengtsson, 1985; Telschow et al., 2002a; Gavrilets, 2004). The fitness graph includes the same information as these two methods and presents a relatively convenient way to calculate reproductive values (Kobayashi and Telschow, 2008). In contrast to the fitness tree, it is a closed graph and eliminates redundancy by preventing reappearance of identical classes. The classes of a class-structured population are linked by arrows (from parent to child class) and a corresponding fitness weight. These denote the net number of offspring of the child class per individual of the parent class, according to genetic contribution. The reproductive value of a class is a weighted sum of the reproductive values of the immediate downstream classes. Thus, the fitness graph is easily converted into a system of coupled equations that has to be solved to attain the reproductive values of all classes. It is important to note that the fitness graph approach requires the system to be at equilibrium in order for the recursive nature of the calculation of reproductive values to be applicable.

We create fitness graphs for three different population scenarios: First, we apply the approach to a homogeneous infected host population and calculate reproductive values of uninfected immigrants. For this scenario, we have to assume perfect *Wolbachia* transmission. We use the term *homogeneous* in the sense that all members of a population have the same cytotype; and *heterogeneous* accordingly means that *Wolbachia*-infected and uninfected hosts coexist. As a second scenario, we analyze the reverse situation of infected immigrants in a homogeneous uninfected population. And finally, we look at reproductive values of infected and uninfected resident and immigrant hosts in a heterogeneous population. In the last two scenarios, we allow for imperfect *Wolbachia* transmission.

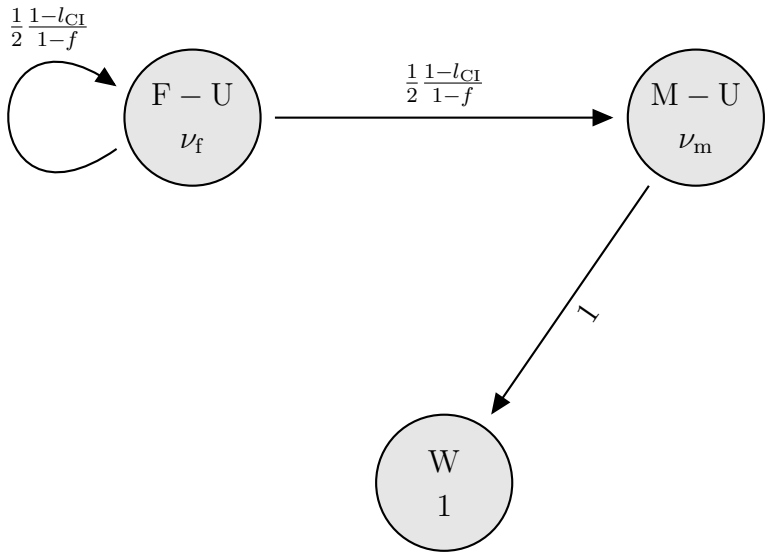


Figure 4.1: Fitness graph for uninfected immigrants in a host population with a *Wolbachia* infection at fixation. Uninfected (U) female (F-U) and male (M-U) immigrants have different reproductive values, ν_{f} and ν_{m}, respectively. Resident hosts are all infected (W) and have a reproductive value of one. An arrow pointing from a parent to a child class and the corresponding fitness weight give the net amount of child-class offspring being produced per parent-class individual. The nominal fecundity deficit of resident females due to their *Wolbachia* infection is transformed into a fecundity benefit (factor $\frac{1}{1-f}$) of immigrant females.

4.2.1 Homogeneous infected population

We start our analysis with migration of uninfected hosts into a host population where the *Wolbachia* infection is at fixation, i.e., a homogeneous infected population. This implies that transmission of the bacteria is perfect, $t = 1$. By definition, the reproductive value of residents is one[1]. However, compared to uninfected immigrant females, resident females suffer a fecundity reduction, $1 - f$. To resolve this incongruity, we have to transform this disadvantage of resident females into a relative advantage of immigrant females. To be precise, we account for the relative fecundity advantage of uninfected female immigrants over infected females by applying a factor $\frac{1}{1-f}$ to the amount of offspring they produce. Note that the ratio of the fecundities of infected and uninfected females remains unchanged by this transformation. In Fig. 4.1, the fitness graph for uninfected immigrants in a homogeneous infected population is shown. It is assumed that immigrants are so rare that they exclusively mate with resident individuals. Because *Wolbachia* transmission is perfect, uninfected females and males are only produced in matings of uninfected females with infected males. In these matings, CI results in a fraction $1 - l_{\mathrm{CI}}$ of fertilized eggs developing properly, and as each sex contributes one half

[1] In homogeneous populations, the per-capita and per-class reproductive values of residents are equal because there is only one class with a frequency of one. The per-class reproductive values of immigrants in a homogeneous population however is zero because their frequency is assumed to be negligible.

of the offspring's genes, a fitness weight of $\frac{1}{2}\,\frac{1-l_{\mathrm{CI}}}{1-f}$ is attained for the production of uninfected females and males. Uninfected males, on the other hand, are fully compatible with resident females, so that they contribute to the production of infected/resident hosts as if they were residents themselves. This gives a fitness weight of one. From the graph, it is evident that immigrant genes can enter the resident gene pool only via males because all offspring of female immigrants are uninfected. It is straightforward to translate the fitness graph into a system of coupled equations for the reproductive values of immigrant females, ν_{f}, and males, ν_{m}:

$$\nu_{\mathrm{f}} = \frac{1}{2}\,\frac{1-l_{\mathrm{CI}}}{1-f}\,(\nu_{\mathrm{f}}+\nu_{\mathrm{m}})\,, \tag{4.5a}$$

$$\nu_{\mathrm{m}} = 1. \tag{4.5b}$$

The system's solutions are also straightforward to calculate and read

$$\nu_{\mathrm{f}} = \frac{1-l_{\mathrm{CI}}}{1-2\,f+l_{\mathrm{CI}}}, \tag{4.6a}$$

$$\nu_{\mathrm{m}} = 1. \tag{4.6b}$$

Assuming an unbiased sex ratio, the average reproductive value of an uninfected immigrant (u) in the *Wolbachia* (w) infected population therefore computes to

$$\bar{\nu}^{\mathrm{u}\to\mathrm{w}} = \nu_{\mathrm{u}} = \frac{1}{2}\,(\nu_{\mathrm{f}}+\nu_{\mathrm{m}}) = \frac{1-f}{1-2\,f+l_{\mathrm{CI}}}. \tag{4.7}$$

Effective migration from an uninfected mainland to an all infected island

In a two-population scenario with migration from an uninfected mainland population to an island population where *Wolbachia* is fixed (see Fig. 3.1a), the effective migration can then be estimated by using Eq. (4.4):

$$m_{\mathrm{eff}}^{\mathrm{u}\to\mathrm{w}}\big|_{t=1} \approx \frac{(1-f)\,m}{1-2\,f+l_{\mathrm{CI}}}, \tag{4.8}$$

if it holds that $l_{\mathrm{CI}} > f$. Otherwise, the infection will not be maintained in the island population (see Section 3.2.5.1), so that effective migration is equal to the real migration. In Fig. 4.2, the gene flow factor ($gff = m_{\mathrm{eff}}/m$) for migration from an uninfected mainland into an infected island population is plotted as a function of the level of CI. It decreases monotonously with stronger cytoplasmic incompatibility because female migrants suffer greater offspring losses. Numerical simulations of the system reveal that the approximation of the gene flow factor by $\bar{\nu}^{\mathrm{u}\to\mathrm{w}}$ yields good results not only if migration is weak but also if CI is strong and migration occurs at considerable rates (for details on how the numerical simulations were conducted, see Section 4.2.5). The gene flow factor monotonously increases with larger fecundity reduction in infected females. This reflects the increasing relative fecundity advantage that uninfected female immigrants have over resident females. If the fecundity costs are such that the infection on the island is barely stable, i.e., $f \lesssim l_{\mathrm{CI}}$ (compare Sections 3.2.1 and 3.2.5.1), it holds that $gff = 1$, and gene flow is not impeded by CI.

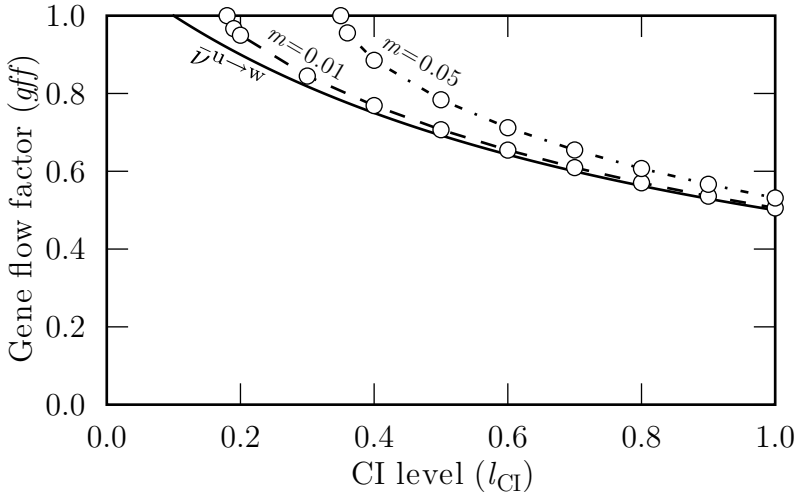


Figure 4.2: Effective migration for uninfected immigrants in a host population with a *Wolbachia* infection at fixation. The gene flow factor, $gff = \frac{m_{\text{eff}}}{m}$, is one if $l_{\text{CI}} \leq f$, and monotonously decreases with larger CI levels. Open circles represent data from numerical simulations for two different migration rates. For weak migration, the average reproductive value of an uninfected immigrant, $\bar{\nu}^{\text{u}\to\text{w}}$ as per Eq. (4.7), can be used to approximate the gene flow factor (solid line). Note that $\bar{\nu}^{\text{u}\to\text{w}}$ still yields good estimations for *gff* even for considerable migration if CI is strong. Parameters: $f = 0.1$, $t = 1$.

We conclude our analysis of migration into a homogeneous infected population by considering two special cases.

No fecundity costs. First, in the case that infected females do not incur fecundity costs, $f = 0$, Eq. (4.8) simplifies to

$$m_{\text{eff}}^{\text{u}\to\text{w}}\big|_{f=0,t=1} \approx \frac{m}{1+l_{\text{CI}}}. \tag{4.9}$$

The stronger the cytoplasmic incompatibility, the more is gene flow into the population effectively reduced. Note that this is the same formula that Telschow et al. (2002a) derived for the scenario with perfect *Wolbachia* transmission and without fecundity reduction.

Complete incompatibility. A second special case occurs if cytoplasmic incompatibility is complete, i.e., l_{CI}=1. Then, the effective migration rate simply reads

$$m_{\text{eff}}^{\text{u}\to\text{w}}\big|_{l_{\text{CI}}=1,t=1} \approx \frac{m}{2}. \tag{4.10}$$

In this case, only genes from immigrant males can enter the resident gene pool because immigrant females lose all their offspring due to CI matings. In the fitness graph (Fig. 4.1), the class of uninfected females is isolated, and only the arrow between uninfected males and residents remains. This also explains why the approximation by Eq. (4.4) yields good results at high CI levels even for relatively large migration rates (see Fig. 4.2).

Maximal gene flow reduction. Eq. (4.10) gives the maximal gene flow reduction that is possible in the present scenario. Gene flow is therefore reduced strongest if both transmission fidelity of *Wolbachia* and the level of CI are close to one. To conclude, a stable *Wolbachia* polymorphism can reduce gene flow from an uninfected mainland population into an infected island population at best to 50%.

4.2.2 Homogeneous uninfected population

As a second scenario, we analyze migration into an uninfected host population. By definition, the average reproductive value of resident members of an uninfected host population is one. Thus, an uninfected immigrant's reproductive value will also be one, and gene flow between uninfected populations will be unimpeded. However, infected immigrants suffer offspring losses due to CI matings of male immigrants and descendants and also due to the fecundity disadvantage of female immigrants and descendants. In Fig. 4.3, the fitness graph for infected female and male immigrants in an uninfected host population is depicted. The fitness weights are straightforward to compute. E.g., infected females contribute $(1-f)(1-t)$ to the production of uninfected hosts. Immigrant's genes can enter the resident gene pool both through females and males. The gene flow is hampered by CI losses on the male side and by fecundity costs (relative to uninfected resident females) and *Wolbachia* transmission to offspring (only genes in uninfected offspring become part of the resident gene pool) on the female side.

From the fitness graph, the following system of equations for the reproductive

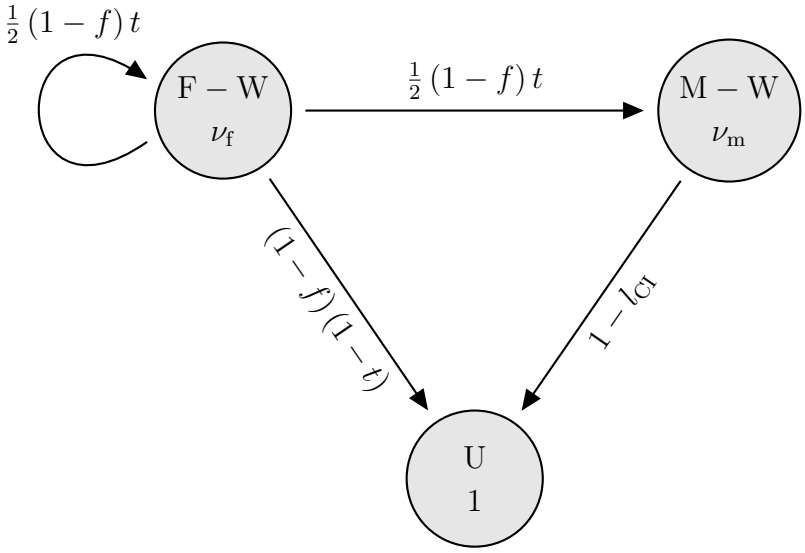


Figure 4.3: Fitness graph for infected immigrants in an uninfected host population. *Wolbachia* (W) infected female (F-W) and male (M-W) immigrants have different reproductive values, ν_{f} and ν_{m}, respectively. Resident hosts are all uninfected (U) and have a reproductive value of one. An arrow pointing from a parent to a child class and the corresponding fitness weight give the net amount of child-class offspring being produced per parent-class individual. E.g., male immigrants only contribute to the production of uninfected hosts, weighed with a factor $1-l_{\mathrm{CI}}$ which reflects the loss of offspring due to CI.

values of female and male immigrants can be built:

$$\nu_{\mathrm{f}} = \frac{1}{2}\,(1-f)\,t\,(\nu_{\mathrm{f}}+\nu_{\mathrm{m}}) + (1-f)\,(1-t)\,, \tag{4.11a}$$

$$\nu_{\mathrm{m}} = 1 - l_{\mathrm{CI}}. \tag{4.11b}$$

Solving Equations 4.11, the reproductive values can be written as functions of the infection's dynamical parameters:

$$\nu_{\mathrm{f}} = \frac{(1-f)\,[2-(1+l_{\mathrm{CI}})\,t]}{2-(1-f)\,t}, \tag{4.12a}$$

$$\nu_{\mathrm{m}} = 1 - l_{\mathrm{CI}}. \tag{4.12b}$$

Then, assuming an unbiased sex ratio, the average reproductive value of an infected immigrant in an uninfected population is

$$\nu_{\mathrm{w}} = \frac{1}{2}\,(\nu_{\mathrm{f}}+\nu_{\mathrm{m}}) = 1 - \frac{f+l_{\mathrm{CI}}}{2-(1-f)\,t}. \tag{4.13}$$

Effective migration from an infected mainland to an uninfected island. To estimate the effective migration in a two-population scenario with an infected mainland and an uninfected island (see Fig. 3.1a), we assume that the *Wolbachia* infection on the mainland is at the stable equilibrium frequency

$$x_{main}(f, l_{\mathrm{CI}}, t) = \frac{f+l_{\mathrm{CI}} \mp \sqrt{(l_{\mathrm{CI}}-f)^2 - 4\,(1-f)^2\,l_{\mathrm{CI}}\;t\,(1-t)}}{2\,l_{\mathrm{CI}}\,[f+(1-f)\,t]}, \tag{4.14}$$

as calculated in Section 3.2.1 of the previous chapter (see Eq. (3.2b)). Hence, the average reproductive value of an immigrant from the *Wolbachia* (w) infected mainland in the uninfected (u) island population is the sum of the reproductive values of immigrants weighed by their frequency on the mainland:

$$\begin{aligned} \bar{\nu}^{\mathrm{w}\to\mathrm{u}} &= x_{main}(f, l_{\mathrm{CI}}, t)\cdot\nu_{\mathrm{w}} + \big(1 - x_{main}(f, l_{\mathrm{CI}}, t)\big)\cdot 1 \\ &= 1 - x_{main}(f, l_{\mathrm{CI}}, t)\;\frac{f+l_{\mathrm{CI}}}{2-(1-f)\,t}. \end{aligned} \tag{4.15}$$

We can then use Eq. (4.15) to derive an approximation of the effective migration rate for the migration from an infected mainland population into an uninfected island population:

$$m_{\mathrm{eff}}^{\mathrm{w}\to\mathrm{u}} \approx m\left(1 - x_{main}(f, l_{\mathrm{CI}}, t)\;\frac{f+l_{\mathrm{CI}}}{2-(1-f)\,t}\right), \tag{4.16}$$

if $l_{\mathrm{CI}} > l_{\mathrm{crit}}(f,t)$, $t > \frac{1}{2}$, and $m < m_{\mathrm{crit}}^{\mathrm{w}\to\mathrm{u}}(f, l_{\mathrm{CI}}, t)$, i.e., if both the infection can be stably maintained on the mainland (see Eq. (3.4)) and migration is below the critical migration rate (see Eq. (3.15)). Otherwise, unidirectional CI does not impede gene flow, and the effective migration is equal to the real migration. Fig. 4.4 shows how the ratio of effective to real migration (the gene flow factor, *gff*), depends on the CI level. If $l_{\mathrm{CI}} > l_{\mathrm{crit}}$ then the gene flow factor monotonously decreases with larger

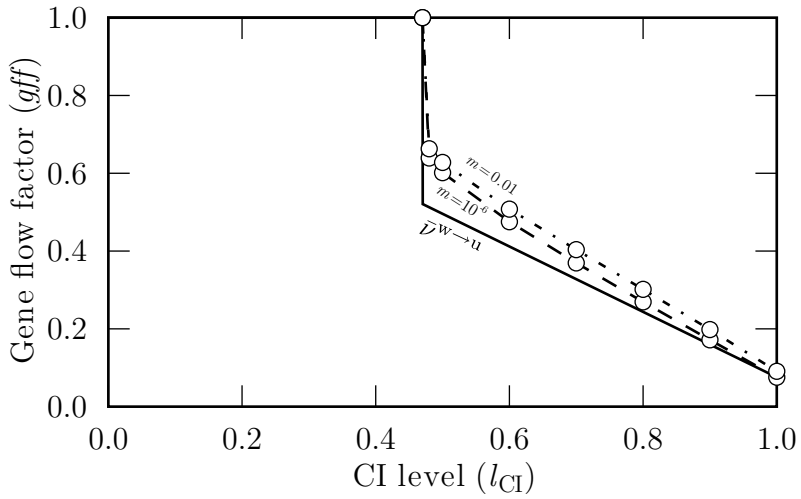


Figure 4.4: Effective migration from a *Wolbachia* infected mainland into an uninfected host population. The gene flow factor, $gff = \frac{m_{\text{eff}}}{m}$, is one if $l_{\text{CI}} \leq l_{\text{crit}}$, and monotonously decreases with larger CI levels. Open circles represent data from numerical simulations for two different migration rates. The average reproductive value of an immigrant from the (mainly) infected mainland, $\bar{\nu}^{\text{w}\to\text{u}}$, can be used to approximate the gene flow factor (solid line). The approximation is best for weak migration and strong CI. Parameters: $f = 0.1$, $t = 0.9$.

CI levels, the slope being steepest for CI levels just above the critical level. Because *Wolbachia* transmission is imperfect, immigrant genes still enter the resident gene pool even if cytoplasmic incompatibility is complete. By means of numerical simulations, we could confirm that Eq. (4.4) yields good approximations for all migration rates below the critical rate, especially if CI effects are strong (for details on how we performed numerical simulations, see Section 4.2.5). With increasing fecundity costs for infected females, the gene flow factor monotonously decreases. The same is true for increasing transmission rates. In both cases, it becomes more and more difficult for genes in female immigrants to enter the resident, uninfected gene pool, whereas genes in male immigrants are unaffected by these parameter changes.

For the scenario of migration from a (mainly) infected mainland population into an uninfected island population, three special cases can be investigated.

Perfect transmission. First, if the fidelity of *Wolbachia* transmission is one, $t = 1$, Eq. (4.16) greatly simplifies because $x_{main}(f, l_{\text{CI}}, 1) = 1$. The effective migration rate in this case is

$$m_{\text{eff}}^{\text{w}\to\text{u}}\big|_{t=1} \approx m\ \frac{1 - l_{\text{CI}}}{1 + f}, \tag{4.17}$$

if $l_{\text{CI}} > f$ and $m < m_{\text{crit}}^{\text{w}\to\text{u}}(f, l_{\text{CI}}, 1) = \left(1 - \sqrt{1 - f}\right)^2 / l_{\text{CI}}$. This formula can serve as a general estimate of gene flow reduction due to unidirectional CI because transmission rates of *Wolbachia* and infection frequencies tend to be close to one in nature. It is easy to see that effective migration is reduced both due to cytoplasmic incompatibility and due to the fecundity costs that infected females incur.

No fecundity costs. As a second special case, we analyze effective migration if *Wolbachia* infected females do not suffer fecundity costs, $f = 0$. Then, Eq. (4.16) reduces to

$$\begin{aligned} m_{\text{eff}}^{\text{w}\to\text{u}}\big|_{f=0} &\approx m\left(1 - \frac{l_{\text{CI}}\, x_{main}(0, l_{\text{CI}}, t)}{2-t}\right), \\ &\approx m\left(1 - \frac{l_{\text{CI}} + \sqrt{l_{\text{CI}}^2 - 4\, l_{\text{CI}}\, t\,(1-t)}}{2\,t\,(2-t)}\right), \end{aligned} \tag{4.18}$$

if $l_{\text{CI}} > 4\,t\,(1-t)$, $t > \frac{1}{2}$, and $m < m_{\text{crit}}^{\text{w}\to\text{u}}(0, l_{\text{CI}}, t)$ (cp. Section 3.2.5.2). Again, it is evident that CI results in a reduction of effective migration. However, that reduction is to some degree counteracted by imperfect *Wolbachia* transmission as immigrant genes "leak" into the uninfected population.

Complete incompatibility. The third special scenario occurs if CI is complete, $l_{\text{CI}} = 1$. Then, the reproductive value of males is zero, and gene flow can occur only via females (cp. the fitness graph in Fig. 4.3). Furthermore, the infection is at fixation on the mainland, $x_{main}(f, 1, t) = 1$, so that Eq. (4.16) greatly simplifies:

$$m_{\text{eff}}^{\text{w}\to\text{u}}\big|_{l_{\text{CI}}=1} \approx m\ \left(1 - \frac{1+f}{2-(1-f)\,t}\right), \tag{4.19}$$

if $t > \frac{1}{2}$ and $m < m_{\text{crit}}^{\text{w}\to\text{u}}(f, 1, t)$. The influence of the two parameters, f and t, as described above in the first two special scenarios, can be seen in Eq. (4.19) as well: Greater fecundity costs and better transmission fidelity both result in a larger gene flow reduction.

We want to remark that our analysis is in line with previous studies. E.g., Telschow et al. (2002b) investigated a scenario with an infected mainland and an uninfected island for the case that transmission is perfect and fecundity costs do not occur, and calculated the effective migration rate to be

$$m_{\text{eff}}^{\text{w}\to\text{u}}\big|_{f=0,t=1} \approx m\,(1 - l_{\text{CI}})\,. \tag{4.20}$$

This formula can be attained by setting $f = 0$ in Eq. (4.17). Furthermore, in Telschow et al. (2007) the constraint of no fecundity costs was waived, and the resulting effective migration rate is correctly reproduced by Eq. (4.17).

Maximal gene flow reduction. Finally, we state that the maximal reduction in gene flow in the present scenario with migration from an infected mainland to an uninfected island is easily found to be zero, $m_{\text{eff}}^{\text{w}\to\text{u}} \approx 0$, for perfect transmission of *Wolbachia* and a complete CI level, $t = 1$ and $l_{\text{CI}} = 1$, respectively. This means that if *Wolbachia* is not able to spread in the island population, *no* genes will make it into the resident gene pool, and reproductive isolation in this direction of migration is complete. Note that for the same set of parameter values, gene flow in the other direction is also maximally reduced albeit only to 50%.

4.2.3 Heterogeneous population

In order to calculate effective migration rates for migration into a heterogeneous host population where infected and uninfected individuals coexist, we first need to calculate the reproductive values of all classes that the population is composed of. There are four classes: uninfected females (which we refer to by the label F-U, or later in matrix and vector notation by the index 1) and males (M-U, 2), as well as infected females (F-W, 3) and males (M-W, 4). In Fig. 4.5, the fitness graph for the four classes is shown. All classes are mutually connected to all other classes (including themselves), with the notable exception of uninfected females who do not contribute to the production of infected females and males. In contrast to the fitness graphs in the previous Sections 4.2.1 and 4.2.2, class production depends not only on the infection dynamical parameters, f, l_{CI}, and t, but also on the infection frequency itself, x. For example, an uninfected male (M-U) contributes to the production of the class of infected males (M-W) when mating with an infected female (F-W) whose fecundity is reduced by the factor $1-f$, and who produce infected offspring at the rate t. The focal male will meet such a female with probability x, and it will contribute half of the offspring's genes. Put together, the corresponding weight is

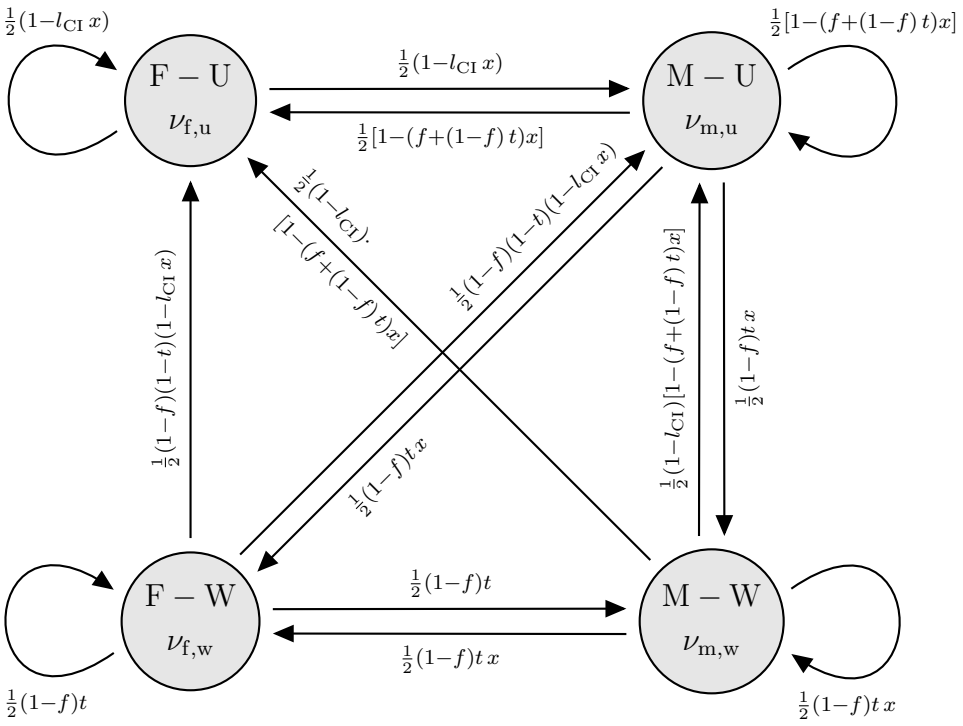


Figure 4.5: Fitness graph for a heterogeneous population. Reproductive values (ν) of *Wolbachia* (W) infected and uninfected (U) female (F) and male (M) hosts depend on the equilibrium frequency of the *Wolbachia* infection, x. An arrow pointing from a parent to a child class and the corresponding weight depict the production of child-class offspring per parent-class host, according to genetic contribution. E.g., an uninfected male genetically contributes $\frac{1}{2}\,(1-f)\,t\,x$ to the production of infected males by mating with an infected female (see main text for more detail on how to calculate the fitness weight).

$\frac{1}{2}(1-f)\,t\,x$. All other fitness weights can be derived accordingly.

For the heterogeneous host population, it is convenient to formulate class dynamics in matrix notation. The equation system for the reproductive values can then be stated in the language of eigenvectors. The fitness graph can be translated into a fitness matrix[2]:

$$L = \frac{1}{2}\begin{pmatrix} 1-l_{\mathrm{CI}}\,x & 1-\left[f+(1-f)\,t\right]x & (1-f)\,(1-t)\cdot(1-l_{\mathrm{CI}}x) & (1-l_{\mathrm{CI}})\left[1-(f+(1-f)\,t)\,x\right] \\ 1-l_{\mathrm{CI}}\,x & 1-\left[f+(1-f)\,t\right]x & (1-f)\,(1-t)\cdot(1-l_{\mathrm{CI}}x) & (1-l_{\mathrm{CI}})\left[1-(f+(1-f)\,t)\,x\right] \\ 0 & (1-f)\,t\,x & (1-f)\,t & (1-f)\,t\,x \\ 0 & (1-f)\,t\,x & (1-f)\,t & (1-f)\,t\,x \end{pmatrix}, \tag{4.21}$$

where the matrix elements, l_{ij}, correspond to the amount of class-i offspring being produced per class-j individual, according to genetic contribution. If we write the frequencies of the four classes as a vector,

$$\vec{X} = \begin{pmatrix} x_{\mathrm{f,u}} \\ x_{\mathrm{m,u}} \\ x_{\mathrm{f,w}} \\ x_{\mathrm{m,w}} \end{pmatrix}, \tag{4.22}$$

then class dynamics can be formulated as $r\vec{X}' = L\vec{X}$ where r is the growth rate of the population and equals the total amount of offspring produced: $r = \sum_{i,j} l_{ij}x_j = \left|L\vec{X}\right|$. At equilibrium, the *Wolbachia* frequency must be constant, $x = x^*$, so assuming an unbiased sex ratio, we can write

$$\vec{X}^* = \frac{1}{2}\begin{pmatrix} 1-x^* \\ 1-x^* \\ x^* \\ x^* \end{pmatrix} \tag{4.23}$$

and

$$L^* = L\big|_{x=x^*}. \tag{4.24}$$

Note that $\vec{X}^*$ is a right eigenvector of the equilibrium fitness matrix, L^*. The eigenvalue r^* characterizes the total amount of offspring at equilibrium.

Eigensystem. Taylor (1990) showed that if eigenvectors on the right of the fitness matrix can be interpreted as frequencies, eigenvectors on the left have an interpretation as future values, i.e., as reproductive values. Thus, if we write the per-capita

[2] In order to guard against confusion with our commonly used abbreviation for *Wolbachia* (W, or w), we deviate from the usual notation for the fitness matrix, i.e., W, and instead denote the fitness matrix with L in reference to the Leslie matrix for age-structured populations (Leslie, 1945).

reproductive values for members of the four classes in vector form,

$$\vec{V} = \begin{pmatrix} \nu_{\mathrm{f,u}} \\ \nu_{\mathrm{m,u}} \\ \nu_{\mathrm{f,w}} \\ \nu_{\mathrm{m,w}} \end{pmatrix}, \tag{4.25}$$

we can set up an eigensystem,

$$r^* \, \vec{V}^* = \vec{V}^* \, L^*, \tag{4.26a}$$

$$1 = \vec{X}^* \, \vec{V}^*, \tag{4.26b}$$

where $\vec{V}^*$ is the equilibrium vector of reproductive values. Eq. (4.26a) represents the mathematical formulation of the fitness graph in Fig. 4.5, and Eq. (4.26b) ensures that the average reproductive value within the population is normalized to one. We can denote the elementwise multiplication operator with $\odot$ (sometimes referred to as the Hadamard product), and write the vector of per-class reproductive values as

$$\vec{C}^* = \begin{pmatrix} c_{\mathrm{f,u}} \\ c_{\mathrm{m,u}} \\ c_{\mathrm{f,w}} \\ c_{\mathrm{m,w}} \end{pmatrix} = \vec{X}^* \odot \vec{V}^*. \tag{4.27}$$

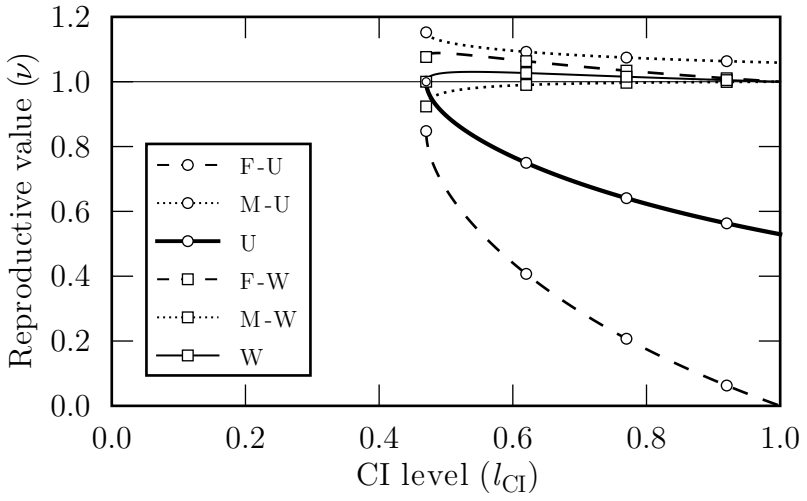
The scalar product in Eq. (4.26b) thus guarantees that the sum over all per-class reproductive values is one so that the interpretation stated in the introductory part of this chapter can be applied, namely that the per-class reproductive value is equal to the probability that a gene (on neutral a marker locus) drawn at random from the population in the distant future originates from an individual of the focal class.

The host population is at equilibrium if the *Wolbachia* frequency is at one of the fixpoint frequencies derived in the previous chapter for a single host population (see Equations (3.2)). For the fixpoint that represents an uninfected population, $x^* = 0$, Equations (4.26) simplify to Equations (4.11), and we have already calculated the solution (see Equations (4.12)). Thus, for the case of a heterogeneous population, we will assume that the infection is at the stable fixpoint frequency

$$x^* = \frac{f + l_{\mathrm{CI}} \mp \sqrt{(l_{\mathrm{CI}} - f)^2 - 4(1-f)^2 l_{\mathrm{CI}}\, t\,(1-t)}}{2\, l_{\mathrm{CI}}\,[f + (1-f)\, t]} \quad \text{as per Eq. (3.2b).} \tag{4.28}$$

Solutions. The solution vector $\vec{V}^* = \vec{V}^*(f, l_{\mathrm{CI}}, t)$ can be calculated analytically. However, it is so cumbersome that we present it only in the Appendix section B. The components of $\vec{V}^*(f, l_{\mathrm{CI}}, t)$ are given by Equations (B.4) and (B.6). The average per-capita reproductive value of an uninfected host, ν_{u}, can be found in Eq. (B.5), and the average reproductive value of an infected individual, ν_{w}, in Eq. (B.7). For complete CI, $l_{\mathrm{CI}} = 1$, these reproductive values correctly simplify to one for infected hosts, and for uninfected hosts to the values that have been derived in Section 4.2.1 (see Equations (4.6) and (4.7)). Also, it can be shown that the average per-class reproductive values of females and males satisfy the demands of Fisher's principle in sex-ratio theory (Fisher, 1930): $c_{\mathrm{f}} = c_{\mathrm{f},u} + c_{\mathrm{f},w} = \frac{1}{2}$, and $c_{\mathrm{m}} = c_{\mathrm{m},u} + c_{\mathrm{m},w} = \frac{1}{2}$.

(a) Per-capita reproductive value



(b) Per-class reproductive value

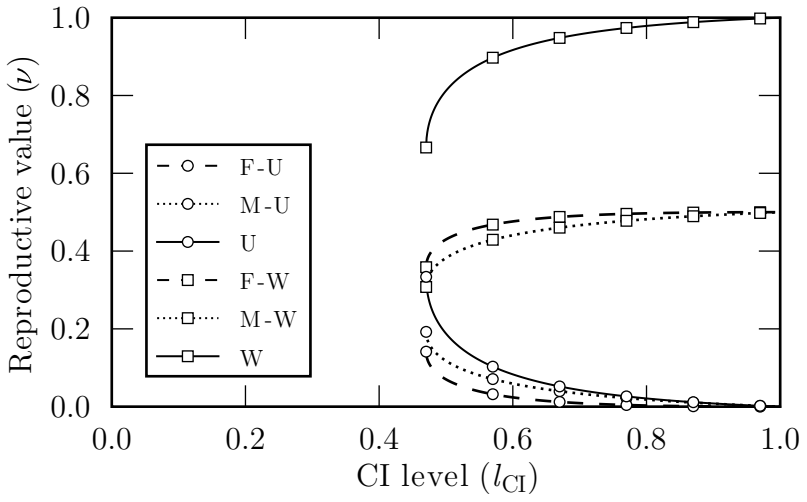


Figure 4.6: Reproductive values in a heterogeneous *Wolbachia* infected population. (a) Individual reproductive values of uninfected females (F-U) and males (M-U), of infected females (F-W) and males (M-W), and of infected (W) and uninfected hosts (U) (assuming an unbiased sex ratio). Note that the latter (emphasized by a heavy line) can be used to estimate the gene flow factor of effective migration from an uninfected mainland into an infected population. (b) Class reproductive values of the four classes by definition add up to one. In both subgraphs, the CI level must be above the critical level, $l_{\mathrm{CI}} > l_{\mathrm{crit}} \approx 0.47$ if the infection is to be maintained in the population. Parameters: $f = 0.1$, $t = 0.9$.

Unfortunately, due to the complexity of the general formulas, it is not easy to derive general conclusions from them. It is however possible to obtain some insight by plotting the reproductive values as functions of the dynamical parameters.

In Fig. 4.6, per-capita (subfigure (a)) and per-class (subfigure (b)) reproductive values are shown as functions of the CI level for fixed fecundity costs and transmission rate, $f = 0.1$ and $t = 0.9$, respectively. The average reproductive value of an individual infected host is larger than one, whereas that of an uninfected host

is smaller than one for all levels of CI, given that CI is strong enough to enable a stable *Wolbachia* infection, $l_{\mathrm{CI}} > l_{\mathrm{crit}} \approx 0.47$. Although infected hosts have a larger reproductive value (per-capita as well as per-class), uninfected hosts are perpetually produced due to imperfect *Wolbachia* transmission. Note however that an uninfected male has the highest per-capita reproductive value of all classes in the population because it is compatible with both classes of females.

Reducing *Wolbachia*-induced fecundity reductions in females, or increasing the transmission fidelity of the bacteria both yield a smaller critical CI level (see Eq. (3.4a)) which in turn results in a stretching of all function graphs in Fig. 4.6. To get an impression of this effect, compare the gene flow factor approximations in Fig. 4.6a (where $t = 0.9$) and in Fig. 4.2 (where $t = 1$). Conversely, increasing fecundity costs, or decreasing the transmission rate result in larger critical CI levels and a compression of the function graphs.

Effective migration from an uninfected mainland to an infected island. We can use the average per-capita reproductive value of an uninfected host to estimate effective migration from an uninfected mainland into an infected (i.e., heterogeneous) island population. By applying Eq. (4.4), and setting the average immigrant reproductive value, $\bar{\nu}^{\mathrm{u}\to\mathrm{w}} = \nu_{\mathrm{u}}$ as per Eq. (B.5), we get

$$m_{\mathrm{eff}}^{\mathrm{u}\to\mathrm{w}} \approx m\,\nu_{\mathrm{u}}. \tag{4.29}$$

The graph of the gene flow factor, $gff(f, l_{\mathrm{CI}}, t)$, as a function of the CI level (at $f = 0.1$ and $t = 0.9$) thus can be taken from Fig. 4.6a (thick solid line with open circles). If the *Wolbachia* infection is barely maintained in the island population, i.e., $l_{\mathrm{CI}} \gtrsim l_{\mathrm{crit}}$, then $gff \lesssim 1$, and gene flow is more or less unimpeded by CI. However, with increasing incompatibility levels, the gene flow factor decreases monotonously until it reaches approximately 0.53 for complete CI. It is remarkable that in a likewise scenario but with migration from a mainland population with a higher infection frequency than on the island (e.g., due to a locally higher *Wolbachia* transmission rate), gene flow would be slightly enhanced because $\nu_{\mathrm{w}} > 1$.

Effective two-way migration between infected and uninfected populations. If migration between an uninfected and an infected population occurs in both directions (see Fig. 3.1c) then the approximations for the mainland-island scenarios can be employed to estimate gene flow between the populations. From Fig. 4.7 where the gene flow factors for both directions of migration are plotted as functions of the real migration rate, m, the asymmetry of effective migration (although real migration is symmetric) is evident. Data from numerical simulations shows that the gene flow factor slightly increases with larger migration rates (for details on how we performed numerical simulations, see the next Section 4.2.5). This is because divergence in cytotype frequencies between the two populations becomes less pronounced. Note that the gene flow factor skips to one as soon as migration exceeds the critical migration rate, $m > m_{\mathrm{crit}}$. Fig. 4.7 illustrates that the analytical gene flow factors derived for the mainland-island scenarios yield good estimations for the gene flow in both directions.

The asymmetry of gene flow reduction due to cytoplasmic incompatibility is best

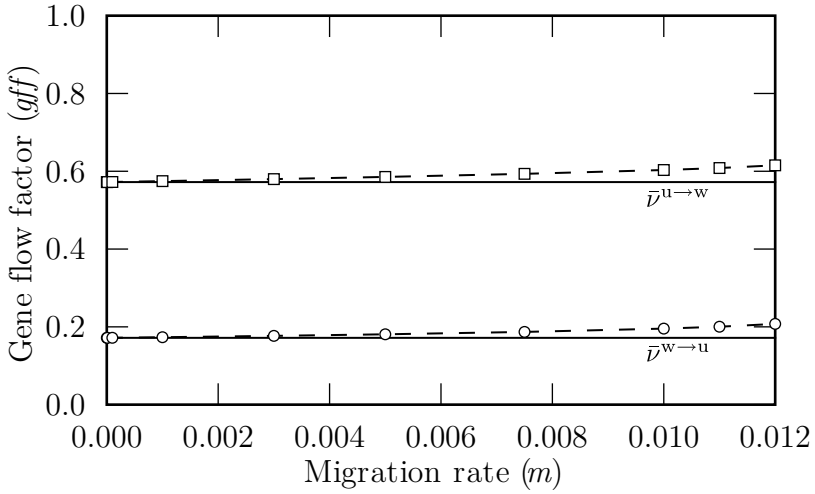


Figure 4.7: Asymmetric gene flow reduction between a *Wolbachia* infected and an uninfected host population. In a scenario with two-way migration, gene flow is reduced in both directions albeit in an asymmetrical manner. The reduction is stronger for migration from the (mainly) infected population into the (mainly) uninfected population (open circles represent data from numerical simulations, and $\bar{\nu}^{w\to u}$ shows the approximation as per Equations (4.15) and (4.16)). In the other direction, i.e., from the uninfected into the infected population, gene flow reduction is less pronounced (numerical data: open squares, approximation: $\bar{\nu}^{u\to w}$ as per Equations (B.5) and (4.29)). In both cases, the quality of the analytical approximation only slightly deteriorates with increasing migration rates. Parameters: $f = 0.1$, $l_{CI} = 0.9$, $t = 0.9$, m symmetric.

understood if we neglect fecundity costs and imperfect transmission, and compare Equations (4.9) and (4.20): $m_{\text{eff}}^{u\to w} \approx m/(1 + l_{CI})$, and $m_{\text{eff}}^{w\to u} \approx m\,(1 - l_{CI})$, respectively. Without CI, the effective migration rate is equal to the real migration rate in both directions. We can easily compute the ratio of the gene flow factors to be $gff^{w\to u}/gff^{u\to w} = 1 - l_{CI}^2$. This immediately shows that for all levels of CI, gene flow from the infected to the uninfected population is reduced more than in the other direction. At $l_{CI} = 1$, i.e., with complete incompatibility, gene flow *out of* the infected population is prevented entirely, and gene flow *into* the infected population is reduced to 50%. Therefore, in a large structure that consists of uninfected and infected patches that are linked by migration, *Wolbachia* turns infected populations into genetic sinks while uninfected populations act as genetic sources.

4.2.4 Local adaptation

In the previous chapter, we showed how local adaptation at a nuclear trait locus, T, of the *Wolbachia* host increases the stability of an infection polymorphism (see Section 3.2.6). Here, we reverse the examination and analyze the effect of CI-induced gene flow reduction on local adaptation. In order to do this, we consider mainland-island scenarios with a stable *Wolbachia* infection divergence. In the island population, trait T_1 is favored by local selection, i.e., individuals harboring T_1 have a $(1 + s)$-times higher viability than individuals showing other traits. The parameter s is called the selection coefficient. On the mainland, selection has lead

to the fixation of a different trait, T_2. First, let us consider a scenario without *Wolbachia*. Then, the frequency dynamics of the locally adaptive allele can be calculated analytically. Let x_{T_1} and x'_{T_1} denote the frequencies of the favored trait on the island, and let m denote the fraction of island hosts that is replaced by migrants from the mainland each generation. Then it holds that

$$x'_{T_1} = \frac{(1-m)\,(1+s)\,x_{T_1}}{1+(1-m)\,s\,x_{T_1}} \equiv F(m,s,x_{T_1})\,. \tag{4.30}$$

In equilibrium, a migration-selection balance will be established. To determine the equilibrium frequency $x^*_{T_1}$ on the island, we solve Eq. 4.30 for $x^*_{T_1} = x'_{T_1} = x_{T_1}$:

$$x^*_{T_1} = 1 - \frac{m}{(1-m)\,s} \equiv G(m,s)\,. \tag{4.31}$$

Next, we consider the general case with a stable infection polymorphism between mainland and island. The dynamics become more complex and cannot be solved analytically anymore (compare Equations (3.52)). However, good analytical approximations can be achieved by using the effective migration rate. Our simple heuristic approach is to substitute m by $m_{\text{eff}}^{\text{w}\to\text{u}}$ in the scenario with an uninfected island (see Eq. (4.16)), and by $m_{\text{eff}}^{\text{u}\to\text{w}}$ in the scenario with an uninfected island (see Eq. (4.29)). A power series expansion of the function G around $m=0$ reveals that for small migration rates, it holds that $x^*_{T_1} \approx G(m_{\text{eff}},s) \approx 1 - \frac{m}{s}\,\bar{\nu}$, where ν is the appropriate average reproductive value of an immigrant in the two scenarios as per Eq. (4.15) or (B.5). Thus, for perfect *Wolbachia* transmission, $t=1$, the following practicable approximations can be derived:

$$x^*_{T_1}(\text{u}\to\text{w}) \;\approx\; 1 - \frac{m}{s}\,\bar{\nu}^{\text{u}\to\text{w}} \;\overset{t=1}{\approx}\; 1 - \frac{m}{s}\,\frac{1-f}{1-2\,f+l_{\text{CI}}} \tag{4.32a}$$

$$x^*_{T_1}(\text{w}\to\text{u}) \;\approx\; 1 - \frac{m}{s}\,\bar{\nu}^{\text{w}\to\text{u}} \;\overset{t=1}{\approx}\; 1 - \frac{m}{s}\,\frac{1-l_{\text{CI}}}{1+f} \tag{4.32b}$$

In Fig. 4.8, the equilibrium frequencies of the locally adaptive trait allele on the island are shown as a function of the real migration rate. The frequency is increased in both mainland-island scenarios, and thereby local adaptations are facilitated. However, because gene flow reduction is asymmetric, so is facilitation of local adaptation. In an uninfected island with migration from an infected mainland, adaptation is better maintained than in an infected island with migration of uninfected hosts. Fig. 4.8 includes both data from numerical simulations and the analytical approximations from Equations (4.32). For details on how we combined modeling of cytoplasmic and nuclear dynamics, and on how we performed numerical simulations, we refer to Section 5.2 of the next chapter.

4.2.5 Numerical simulations

Kobayashi and Telschow (2008) demonstrated how effective migration rates can be numerically calculated. Consider a two-population model of a haploid species reproducing in discrete, non-overlapping generations. Let genetic drift be negligible, and let migration and reproduction take place in this order. Migration is modeled by replacing a fraction m_i in population i with migrants from the respective

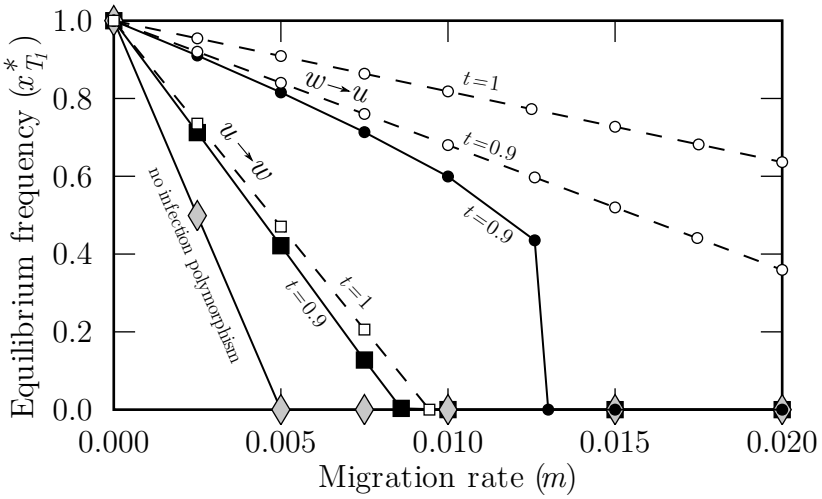


Figure 4.8: Facilitation of local adaptation. In a mainland-island scenario, the island equilibrium frequency of a locally adaptive trait, T_1, is determined by a selection-migration balance. Without *Wolbachia*, the frequencies can be calculated analytically (indicated by gray diamond symbols; see Eq. (4.31)). In comparison, equilibrium frequencies of T_1 are higher and thereby local adaptation is facilitated in scenarios with a stable *Wolbachia* polymorphism. The effect is stronger for an uninfected island with migration of infected hosts (circles; w→u) than for an infected island with migration of uninfecteds (squares; u→w). Filled black symbols show numerically computed equilibrium frequencies, open symbols give analytical approximations according to Equations (4.32). Note that numerical simulations were run with $t = 0.9$, and that even the simplified approximations with $t = 1$ give reasonable estimates. Parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $s = 0.005$, $t = 0.9$.

other population. A "neutral" model, where migrants and residents have the same reproductive success because there are no genetic barriers, serves as a baseline to determine gene flow in models with modifications. In order to define the effective migration rate, frequency dynamics of an allele on a selectively neutral marker locus are considered. If $p_1(\tau)$ and $p_2(\tau)$ denote the frequencies of the marker allele in populations 1 and 2 at time τ, then the frequency dynamics are

$$p_1(\tau+1) = (1-m_1)p_1(\tau) + m_1\, p_2(\tau), \tag{4.33a}$$
$$p_2(\tau+1) = (1-m_2)p_2(\tau) + m_2\, p_1(\tau). \tag{4.33b}$$

It is straightforward to solve equation system (4.33) for the migration rates. This shows that migration rates can be computed if allele frequencies are known. In a system with gene flow modification, the effective migration rates can then be defined as these migration rates taken at the limit $\tau \to \infty$:

$$m_{1,\mathrm{eff}} = \lim_{\tau\to\infty} \frac{p_1(\tau+1) - p_1(\tau)}{p_2(\tau) - p_1(\tau)}, \tag{4.34a}$$
$$m_{2,\mathrm{eff}} = \lim_{\tau\to\infty} \frac{p_2(\tau+1) - p_2(\tau)}{p_1(\tau) - p_2(\tau)}. \tag{4.34b}$$

The effective migration rates are the migration rates that are calculated from ob-

served allele frequencies *as if* the populations had no barriers to gene flow. In a system without gene flow modification, the effective migration rates will correctly yield the real migration rates. Note that the effective migration rates can be calculated from allele frequencies alone (independent of the kind of genetic barriers) which qualifies the approach for numerical analysis and comparison of class-structured systems with various barriers to gene flow.

Numerical calculations of effective migration rates for the systems investigated in this chapter were performed as follows (for the actual equations that model system dynamics, see Chapter 5). The CI system was allowed to reach an equilibrium state which was assumed to have happened if *Wolbachia* frequencies did not change by more than 10^{-7} per generation. Then, a neutral nuclear locus with two alleles was introduced where one allele, N_1, was fixed in population 1, and the other allele, N_2, was fixed in population 2. System dynamics thus involved four classes in each population —uninfected and infected hosts carrying alleles N_1 or N_2— which produced a system of eight coupled difference equations. Nuclear inheritance was independent from the cytoplasmic *Wolbachia* dynamics. The system was numerically simulated for 100 generations, and the frequencies of allele N_1 in generations 99 and 100 were inserted into Equations (4.34) to calculate effective migration rates.

4.3 Discussion

In this chapter, we investigated the impact of *Wolbachia*-induced unidirectional CI on gene flow between host populations of divergent infection states, i.e., between (mostly) uninfected and (mostly) infected populations. We demonstrated both analytically by using a fitness graph approach, and by numerical simulations that gene flow is reduced in both directions in an asymmetric manner. Gene flow is impeded stronger for migration from infected to uninfected populations than in the opposite direction. In larger population structures, infected populations are turned into genetic sinks whereas uninfected populations function as genetic sources. One example for this is that local adaptations are facilitated more in uninfected patches than in infected ones. Our formulas for effective migration can be considered extensions of previous findings by Telschow et al. (2002a), now including the full set of parameters that typically describe *Wolbachia* infection dynamics, and building on a more solid mathematical basis than previously.

Our findings generally support the view that *Wolbachia*-induced unidirectional CI could be a factor in host divergence and speciation. The role of *Wolbachia* would be to reduce gene flow between populations, allowing genetic divergence for locally adaptive traits, and to select for premating isolation (see Chapter 5). Consider a central population with peripheral isolate populations (or islands; see Fig. 4.9). Then, recurrent establishment of peripheral islands from the central source population could produce the circumstances for repeated opportunities for local adaptation and speciation. Our results suggest that peripheral island populations that have lost their *Wolbachia* infection are able to maintain local adaptation in the face of migration better than peripheral populations that maintain their infection. In turn, the threshold migration rate for maintenance of the infection difference is increased, and premating isolation more readily evolves. This means that if recurrent peripheral populations occur, it is the ones that lose their *Wolbachia* that are more likely

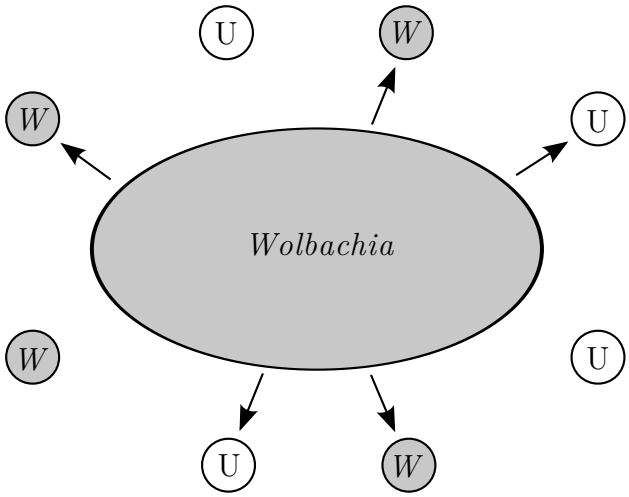


Figure 4.9: Large central population with peripheral islands. The models analyzed in this and the last chapter can be applied to a situation with a large central population and small peripheral islands. The peripheral islands might face migration from the *Wolbachia* (W) infected mainland or not. Some of the islands are infected with *Wolbachia* while others have lost the infection (U). Our results suggest that it is the ones that lose their infection that are better able to maintain local adaptations in the face of migration and that are more likely to diverge into new species.

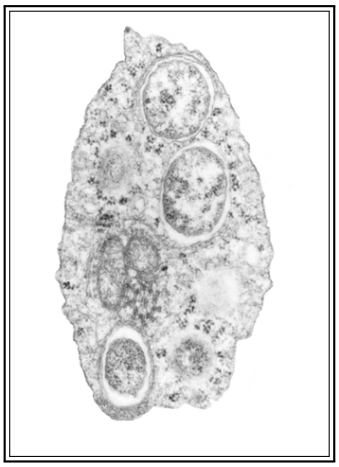
to diverge into new species. This could be considered a form of population selection, where populations that lose their infections are better able to "resist" gene flow and therefore locally adapt and evolve into new species. The scenario is not simply hypothetical. North American populations of the fire ant *Solenopsis invicta*, which were established by presumably small founding populations due to human transport, are devoid of *Wolbachia*, whereas South American source populations show infection polymorphisms (Shoemaker et al., 2000, 2003). Similarly, Atlantic coast populations of the beetle *Chelymorpha alternans* are infected with two *Wolbachia* strains, whereas some Pacific coast populations have lost one of the *Wolbachia* (Keller et al., 2004). Thus, the scenarios envisioned here could occur in nature. A counter example is the spread of *Wolbachia* in uninfected populations of *Drosophila simulans* in uninfected North American populations (Turelli and Hoffmann, 1991).

Our results also have important ramifications for classical reinforcement theory. These will be discussed in more detail in Chapter 5. But a standard critique in reinforcement is that gene flow might overwhelm the selective advantage of being choosy (Mayr, 1963). Although previous work has shown that in principle reinforcement is possible (e.g., Spencer et al., 1986; Liou and Price, 1994; Servedio, 2000), most of the studies are purely simulation based (but see Kirkpatrick and Servedio, 1999). In the present chapter, we developed a method to analyze the effect of unidirectional CI on gene flow analytically. In the next chapter, we combine this approach of gauging gene flow via effective migration rates with modeling the advantage of being choosy by an effective selection coefficient. This allows a deeper understanding of *Wolbachia*-based reinforcement scenarios than with purely simulational methods.

We have shown in Chapter 3 that locally adapted alleles substantially stabilize postzygotic isolation. Other factors that will likely stabilize infection differences and postzygotic isolation are cytoplasmic sex ratio distorters (Engelstädter et al., 2004), nuclear-based mating incompatibilities (Hilgenböcker, 2009, Telschow, priv. comm.), and the accumulation of multiple adaptive genes within the diverging populations. The latter would be particularly important when epistatic interactions among adapted loci occur. As has been shown for a single adapted locus case and for pairwise genetic incompatibilities (Telschow et al., 2002a,b), these effects further enhance the association of infection type with adapted loci, increasing the frequencies of each in the presence of immigration. Another often underrated aspect of local adaptations is that these can directly result in gene flow reduction due to immigrant inviability in foreign habitats (Nosil et al., 2005). This implies that unidirectional CI can —especially in concert with local host adaptation— significantly reduce gene flow even if migration occurs at considerable rates.

A common criticism of the possible role of unidirectional CI in divergence and speciation is that nuclear gene flow is reduced or prevented only in one direction (from the infected to the uninfected population), whereas nuclear genes flow freely in the other direction (e.g., Bordenstein, 2003). In this chapter, we have demonstrated that effective migration from uninfected to infected host populations is significantly reduced by CI, and we have confirmed the previously found general estimation of the gene flow factor, $gff^{\mathrm{u}\to\mathrm{w}} \approx 1/(1+l_{\mathrm{CI}})$ (Telschow et al., 2002a). This means that gene flow reduction is indeed not as pronounced as for migration from the infected to the uninfected population which can be approximated by $gff^{\mathrm{w}\to\mathrm{u}} \approx 1-l_{\mathrm{CI}}$ (see Telschow et al., 2002a, and this chapter). Nevertheless, as argued above stable patterns of unidirectional CI can significantly add to gene flow reduction due to other postzygotic isolation barriers in both directions of migration. Furthermore, the asymmetric gene flow reduction between the populations affects the pattern of adaptation in the *Wolbachia* host; adaptation is favored in the uninfected population but impeded in the infected population. As mentioned above, *Wolbachia* turns infected patches into population genetic sinks, and uninfected into population genetic sources.

In summary, our results demonstrate that if the coexistence of infected and uninfected host populations is stable then unidirectional CI acts as a postzygotic isolation mechanism, reduces gene flow asymmetrically, and selects for local adaptations predominantly in uninfected host populations. This is consistent with the notion that unidirectional CI could be a general factor in speciation processes of arthropods.



Chapter 5

Fisherian runaways and reinforcement of female mating preference

In this chapter, we investigate whether unidirectional cytoplasmic incompatibility selects for premating reproductive isolation between host populations connected by migration. Our analysis of a mathematical model shows that if the infection polymorphism is stable, unidirectional cytoplasmic incompatibility readily selects for premating isolation through spread of female mating preferences in uninfected island populations if they receive migration from a *Wolbachia*-infected mainland. We show that in scenarios with two-way migration, female preferences must be costly in order for reinforcement of premating isolation to evolve. Otherwise, female preferences spread in both the uninfected and infected population which results in Fisherian runaway selection of mating preference and preferred trait, thus genetically homogenizing the populations and consequently leading to increased gene flow. Reinforcement, however, substantially adds to gene flow reduction.

We present data from numerical simulations and analytical approximations for some of our findings. These generally suggest that *Wolbachia*-induced unidirectional cytoplasmic incompatibility can be a factor in divergence and speciation of hosts. In part, the results on reinforcement in mainland-island scenarios have been published (Telschow et al., 2007).

> The grossest blunder in sexual preference, which we can conceive of an animal making, would be to mate with a species different from its own and with which the hybrids are either infertile or, through the mixture of instincts and other attributes appropriate to different courses of life, at so serious a disadvantage as to leave no descendants.
>
> Fisher (1930, p. 130)

5.1 Introduction

Cytoplasmic incompatibility (CI) has attracted attention as a possible mechanism for rapid speciation (Werren, 1997, 1998; Bordenstein, 2003). The basic idea is that CI reduces gene flow between populations, permitting genetic divergence and selecting for premating isolation.

In the case of bidirectional CI there is both empirical and theoretical evidence supporting this view. Field studies show that many insect species harbor different strains of *Wolbachia*, often in different geographic regions (e.g., Merçot et al., 1995). Further, crossing experiments suggest that bidirectional CI is a major isolation factor between some strains and closely related species (e.g., Laven, 1959; Bordenstein et al., 2001). Theoretically, it has been shown that two *Wolbachia* strains can stably coexist in parapatric host populations in the face of substantial migration (Telschow et al., 2005b), and that bidirectional CI reduces the gene flow of locally adapted alleles and selects for premating isolation even if the transmission of *Wolbachia* and the level of incompatibility are incomplete (Telschow et al., 2002a,b; Telschow, 2003; Telschow et al., 2005a). However, it has been argued that bidirectional CI can hardly be considered ubiquitous in nature thus denying a general role in speciation (Bordenstein, 2003; Coyne and Orr, 2004).

Unidirectional CI, on the other hand, is likely to be more common in nature, since it requires that only one population be infected with Wolbachia. However, unidirectional CI is generally not believed to promote speciation in hosts because maintenance of infected and uninfected populations is expected to be unstable in the presence of migration, and because gene flow is presumably reduced in only one direction of hybridization (Coyne and Orr, 2004). We have already addressed these two issues in the previous Chapters 3 and 4, respectively. Our results implicate that *Wolbachia* can add to reproductive isolation if combined with other barriers.

There is, however, an empirical study that indicates that *Wolbachia* could play a larger role by reinforcing premating isolation. In mushroom-feeding *Drosophila* flies in North America, patterns of mating preferences are consistent with reinforcement as an evolutionary response to a unidirectional CI barrier (Jaenike et al., 2006): Uninfected females of *D. subquinaria* that are sympatric with infected *D. recens* show much stronger reluctance to hybrid matings than either *D. recens* females or allopatric *D. subquinaria* females do. Thus, further modeling of the dynamics of unidirectional CI and its possible role in promoting reproductive isolation is needed.

The basic question we address in this chapter is whether there are conditions under which unidirectional CI that occurs in stable infection patterns selects for

premating isolation. We combine the model of *Wolbachia* dynamics of the previous chapters with a well-studied reinforcement model (Servedio, 2000). This new model allows us to investigate the effect of unidirectional CI on genetic divergence of the host. We first consider selection acting on a small (initially uninfected) island population experiencing migration from a large (infected) mainland population, and then extend the model to incorporate two-way migration. The model includes a mating preference locus, a male trait locus undergoing divergent selection in the two populations, and cytoplasmic incompatibility. In a similar effort, Telschow (2003) showed that in such a scenario, unidirectional CI can favor the evolution of female mating preferences. However, due to a runaway process[1] this does not result in premating isolation. In the present study, we extend the reinforcement model to allow for female preferences to be costly.

In general, we present data from numerical simulations of the mathematical model, but in some cases we provide analytical heuristics for a more intuitive understanding of the system. Our results show that premating isolation readily evolves in the considered mainland-island scenario (published in Telschow et al., 2007). In the case of two-way migration, we demonstrate that costs of mating preference are essential if reinforcement processes are to take place.

5.2 Model

We first give a verbal account of the model we use, followed by a mathematical formalization.

5.2.1 Verbal description

Our model is similar to Dobzhansky's (1940) classical model of speciation by reinforcement (see Fig. 5.1). We assume that an ancestral host population has split into two populations. These might be of very different size, like a large mainland and a small island, or of comparable size, like two island populations. The former scenario is analogous to a large central population with a small peripheral isolated population (compare Fig. 4.9). The populations remain for some time in allopatry and diverge during that time at a locus which controls a male trait that might be used in female mate choice. Further, one of the populations is infected with *Wolbachia* but there is no infection in the other population. This might either occur when the ancestral population was uninfected and one of the split populations acquires a novel *Wolbachia* infection, or when the ancestral population was infected and one of the split populations loses its infection. After the establishment of these genetic and cytoplasmic differences, the populations restore contact via migration (secondary contact). In the case of populations of very different size, migration only occurs from the mainland to the island, but if the populations are of comparable size, migration takes place in both directions. For low migration rates and if either *Wolbachia* transmission is incomplete or the infection reduces female fecundity, this infection pattern of infected mainland and (mostly) uninfected island is stable (see

[1] Telschow (2003) called this the "catalyst effect" of *Wolbachia*, because the spread of the mating preference or of the preferred trait through both populations is "catalyzed" by the stable CI-pattern without change of the populations' infection statuses (compare Section 5.3.2).

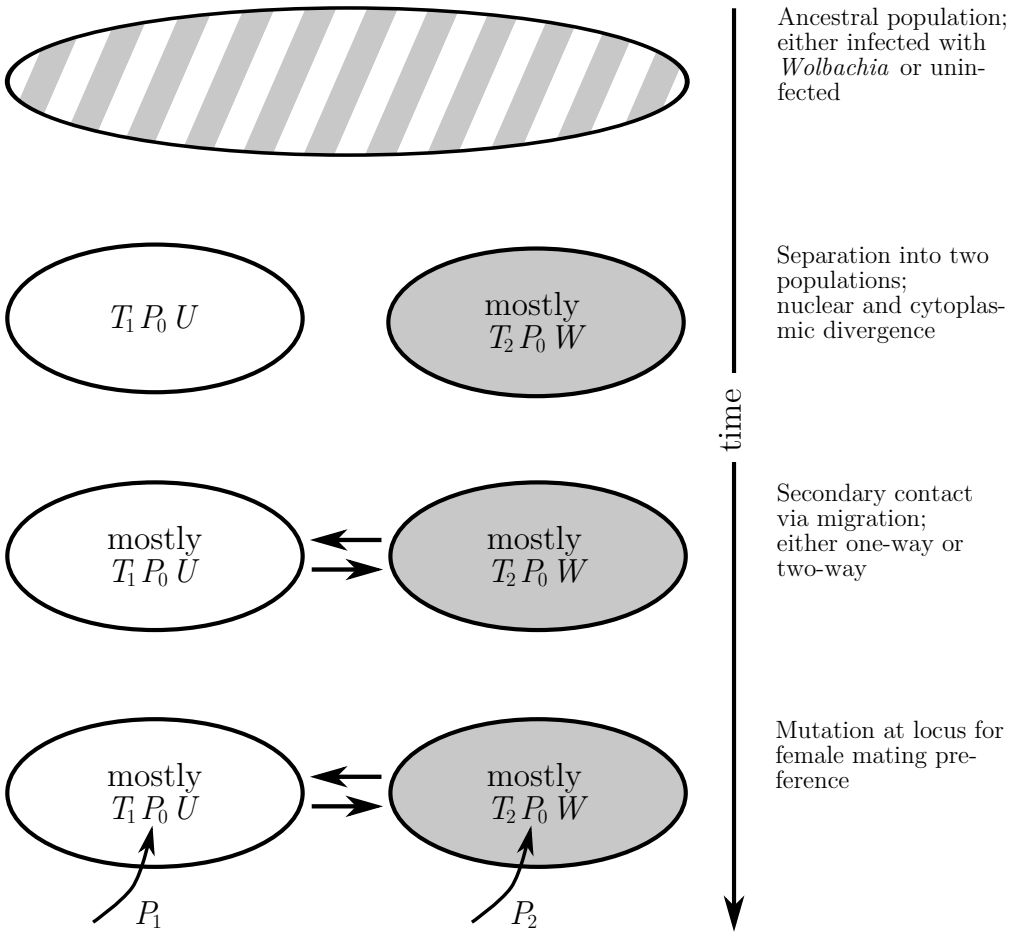


Figure 5.1: Population history. The scenario involves four stages of population history. An ancestral uninfected (or infected with *Wolbachia*) population splits into two separate populations. These populations then diverge at the nuclear trait locus T. Additionally, one population acquires a novel *Wolbachia* infection (or loses its infection). After secondary contact is restored, either by one-way or two-way migration, the infection polymorphism might prove stable, and mutations at a locus that codes for female mating preference occur. For more detail, see main text.

Chapter 3) and, further, unidirectional CI acts as a postzygotic isolation mechanism (see Chapter 4). We investigate whether postzygotic isolation selects for premating isolation and thus reinforces the genetic differences between the populations. To analyze when reinforcement takes place, we introduce new alleles at a locus for female mating preference and study under which circumstances these alleles spread, and whether this results in divergence at the preference locus. Note that the new allele at the preference locus can be interpreted as either a new mutation or as one of various alleles that are always present at low frequencies in a genetically diverse population. In the former case, the allele will appear only in one population, whereas in the latter case it would be introduced in all populations simultaneously.

For simplicity, we assume haploid sexual organisms, an assumption that often has

been made for theoretical analyses involving multiple interacting loci (e.g., Servedio, 2000; Hartl and Clark, 2007). Generations are discrete and non-overlapping. Individuals reproduce sexually with a primary sex ratio of 1:1. The life cycle consists of four steps: migration, viability selection, sexual selection, and reproduction. It is assumed that the first three steps happen in the haploid phase of the organism. The diploid phase occurs during reproduction and ends with the production of the male and female gametes.

Organisms are characterized by their nucleocytotype. Individuals can be either infected with *Wolbachia* or not. Two nuclear loci are considered, a locus P for female mating preference, and a locus T for a male trait. Recombination between the loci may vary from no to full recombination.

The life cycle starts with migration of the haploid organisms between populations 1 and 2. As stated above, in a mainland-island scenario migration occurs only from the mainland to the island. However, in a two-island scenario hosts migrate in both directions. After migration, viability selection takes place. We assume that selection on male trait alleles acts differently in both populations. In population 1, trait T_1 is favored relative to trait T_2. In the other population, selective forces are reversed, and T_2-individuals have a higher viability than T_1-individuals. Therefore, there is divergent selection acting on the trait locus in the two populations.

Mating is nonrandom due to female mating preferences. We assume that males court females indiscriminately, and that a female accepts courtship attempts based on her genotype and the genotype of the courting male. Thereby, females of genotype P_0, P_1, and P_2 have characteristic mating preference strengths. P_0-females are non-discriminating and thus show no mating preference; they are considered to be the wildtype. In contrast, females harboring allele P_1 preferably mate with T_1-males, whereas allele P_2 make females prefer to mate with T_2-males. Being choosy, however, might reduce the overall number of matings a female can have. This happens, for instance, when rejecting an unwanted male increases the risk of predation or loss of time prevents finding another mating partner. We will speak of *costly mating preference*, referring to the reduced number of matings compared to undiscriminating females as costs. The structure of using divergent selection at a trait locus and a preference for that trait is similar to other modeling efforts of mate preference (see Servedio, 2000; Telschow et al., 2005a).

After sexual selection, the haploid individuals produce gametes that fuse to a diploid zygote. In this diploid phase, some individuals suffer from *Wolbachia*-induced unidirectional CI. Following Fine (1978) and our efforts of the last two chapters (see especially Section 3.2), we describe *Wolbachia* dynamics by three parameters: (i) the level of cytoplasmic incompatibility, l_{CI}, (ii) the fecundity reduction, f, of infected females relative to uninfected females, and (iii) the transmission rate of *Wolbachia*, t. Tab. 5.1 gives an overview over the parameters and symbols used in this chapter.

The verbal description of the life cycle can be formalized. The corresponding mathematical model consists of a system of 24 coupled difference equations: 2 populations $\times$ 2 traits $\times$ 3 preference alleles $\times$ 2 cytotypes. A precise mathematical description follows below. In general, the model is too complex to be solved analytically. Therefore, computer simulations were performed to analyze the general model (see Section 5.2.3 for details concerning the simulations), and analytical results are

only given for special cases.

5.2.2 Mathematical formalization

The model verbally described above can be mathematically formalized, the parameters and symbols used are summarized in Tab. 5.1. For a discussion on biologically reasonable parameter values, see Section 3.2. Two nuclear host loci are considered, one locus for female mating preference with the three possible alleles P_0, P_1, and P_2, and one trait locus with the alleles T_1 and T_2. We consider two host populations 1 and 2 that are linked by migration (see Fig. 5.1). Organisms have a life cycle consisting (in that order) of migration, viability selection, sexual selection, and reproduction. The first three steps are assumed to happen in the haploid phase of the organisms. The reproduction step includes the formation of a diploid zygote and the production of female and male gametes with a primary sex ratio of 1:1. There is no overlap between generations.

Individuals are characterized by both their nuclear genotype and cytotype. There are two cytotypes, C_i, or infection statuses: infected with *Wolbachia*, denoted by $C_1 = W$; and uninfected, denoted by $C_2 = U$. The six possible nuclear genotypes are (P_0T_1, P_0T_2, P_1T_1, P_1T_2, P_2T_1, P_2T_2). Overall, this gives twelve different nucleocytotypes. Because our model includes two host populations, the intergenerational change of these nucleocytotype frequencies is described by a system of 24 coupled difference equations. We denote by $x_{i,j,k,l}$ the frequency of nucleocytotype $(T_jP_kC_l)$ in population i, where $i = 1$ is the left population in Fig. 5.1, and $i = 2$ is the right population. Note that we deliberately chose to let all indexes start with one, except for the mating preference index. This may seem inconsistent but it allows for a more intuitive grasping of the link between mating preference and preferred traits (see below), e.g., we can say that P_1-females prefer T_1-males (note the indexes being equal). In order to calculate $x'_{i,j,k,l}$, the frequencies of the nucleocytotypes in the next generation, we take into account the effects of migration, viability selection, sexual selection, and cytoplasmic incompatibility. Thereby, we assume a starting condition where all organisms of one population have the same nucleocytotype, usually T_1P_0U in the left population, and T_2P_0W in the right population.

The intergenerational transition is split into four steps: Migration, viability selection, sexual selection, and reproduction.

Migration

The first step in the life cycle of the hosts is the migration of the haploid individuals. Thereby a fraction of each population is replaced by individuals from the respective other population. We denote this fraction of migrants into population i by m_i, the migration rate. It holds that

$$x^+_{1,j,k,l} = (1 - m_1)\, x_{1,j,k,l} + m_1\, x_{2,j,k,l}, \tag{5.1a}$$

$$x^+_{2,j,k,l} = (1 - m_2)\, x_{2,j,k,l} + m_2\, x_{1,j,k,l}, \tag{5.1b}$$

where $x^+_{i,j,k,l}$ are the frequencies of nucleocytotype $T_jP_kC_l$ in population i after migration. In a mainland-island scenario, one of the two migration rates is assumed to

be zero. With migration occurring in both directions, we will usually only consider symmetrical migration, $m_1 = m_2 \equiv m$.

Viability selection

The second step in the life cycle is the selection acting on the trait locus T. In the left population 1, T_1 is favored by a factor of $1+s$ before T_2. In the right population

Symbol	Description
a_i	Strength of mating preference allele P_i—females harboring P_i mate with T_i-males $1 + a_i$ times more often than undiscriminating females do; can be transformed into rejection probability p_i; $0 \leq a_i < \infty$, $i \in \{0, 1, 2\}$
f	Fecundity costs of a *Wolbachia* infection—infected females' egg production is reduced by f; $0 \leq f < 1$
l_{CI}	Level of cytoplasmic incompatibility (CI level)—fraction of zygotes that die if an uninfected egg is fertilized by sperm from an infected male; $0 \leq l_{\mathrm{CI}} \leq 1$
m_i	Migration rate—fraction of a host population i that is replaced by migrants from the respective other population; $0 \leq m_i \leq 1$, $i \in \{1, 2\}$
p_i	Rejection probability of mating preference allele P_i—females carrying P_i reject non-T_i males with probability p_i; can be transformed into mating preference strength a_i; $0 \leq p_i \leq 1$, $i \in \{0, 1, 2\}$
q	Transition probability—probability of another mating round after rejection of an unwanted male; $0 \leq q \leq 1$
r	Recombination rate—recombination may occur between nuclear loci T and P at rate r; $0 \leq r \leq \frac{1}{2}$
s	Selection coefficient—hosts carrying a locally adaptive allele at the nuclear locus T have increased viability by factor $1+s$; $s \geq 0$
t	Transmission rate—fraction of an infected females' eggs that inherit *Wolbachia*; $0 < t \leq 1$
$x_{i,j,k,l}$	Frequency of hosts in population i with trait allele T_j, preference allele P_k, and cytotype C_l; $i \in \{1, 2\}$, $j \in \{1, 2\}$, $k \in \{0, 1, 2\}$, $l \in \{1, 2\}$
x_W	*Wolbachia* infection frequency within a host population
C_i	Cytotype—infection status; $i \in \{1, 2\}$ where $C_1 = W$ denotes infected hosts, $C_2 = U$ uninfected ones
P_i	Allele at host nuclear locus P for female mating preference; $i \in \{0, 1, 2\}$
T_i	Allele at host nuclear trait locus T; $i \in \{1, 2\}$

Table 5.1: Glossary of model notation. Overview of the parameters and symbols used in the model of this chapter. For more details, see main text.

2 however, individuals with genotype T_2 have a $(1+s)$-times higher fitness than individuals with genotype T_1. Viability selection at the T locus is described by

$$x^{++}_{i,j,k,l} = \begin{cases} \frac{1+s}{W_i}\, x^{+}_{i,j,k,l} & \text{if } i=j \\ \frac{1}{W_i}\, x^{+}_{i,j,k,l} & \text{else,} \end{cases} \tag{5.2}$$

where W_i is the sum over indexes j, k, and l of all numerators in Eq. (5.2) for population i. $x^{++}_{i,j,k,l}$ are the frequencies of nucleocytotype $T_jP_kC_l$ in population i after viability selection. In a mainland-island scenario, viability selection will result in fixation of the adaptive trait on the mainland.

Sexual selection

To model sexual selection we build on concepts that were originally developed by Kirkpatrick (1982) for a model of runaway sexual selection, that were extended by Servedio and Kirkpatrick (1997) and Servedio (2000) to study reinforcement scenarios based on nuclear incompatibilities, and that were adopted to the case of CI by Telschow et al. (2005a). Thereby, we introduce a weighting factor $S_{i,j,k}$ that reflects how mating frequencies are affected by female mating preferences. These are a measure of how often females of genotype P_j mate with males of genotype T_k in population i. The parameter a_j indicates the mating preference strength of allele P_j, and has a range $0 \leq a_j \leq \infty$. Wildtype females of genotype P_0 show no mating preference, and $a_0 = 0$. Let $y^{++}_{i,j}$ be the frequency of allele T_j in population i after viability selection, then the weighting factor can be defined as

$$S_{i,j,k} = \begin{cases} 1 & \text{if } j=0 \\ \dfrac{1+a_j}{1-(1-a_j)\, y^{++}_{i,j}} & \text{if } j=k \\ \dfrac{1}{1-(1-a_j)\, y^{++}_{i,j}} & \text{if } j>0 \ \wedge \ j\neq k. \end{cases} \tag{5.3}$$

P_1-females mate $1+a_1$ times as often with T_1-males as the undiscriminating females of genotype P_0 do. The same holds true for P_2-females and T_2-males. The denominator in Eq. (5.3) guarantees that choosy females have the same number of matings as wildtype females, lest they gain a benefit by simply acquiring more matings. Female mating preference thus is "free" by definition, there are no costs involved.

In this study however, we also aim at investigating costly mating preference. In order to model such costs, we introduce two probability parameters. The first one is the *rejection probability* p_i and describes the probability by which a choosy female of genotype P_i (where $i \in \{1,2\}$) rejects any unwanted male, i.e., any male not showing trait T_i. For wildtype P_0-females, the according rejection probability is assumed to be $p_0 = 0$, and all males they encounter are accepted as mating partners. In contrast, females of genotype P_1 will accept any T_1-male but reject T_2-males with probability p_1. For P_2-females, this mating preference is reversed. In our model, we assume that $0 < p_i \leq 1$ for $i \in \{1,2\}$. The second parameter we introduce, q, gives the probability that a female gets another chance to find a mating partner

after having rejected a male. In other words, it is the probability of transition into a new mating round before the life cycle moves to the reproduction step. Therefore, we call q the *transition probability*, and we assume that $0 \leq q \leq 1$. Thus, mating rounds are treated as a simple Markov chain (Markov, 1971).

The weight that models costly sexual selection can be derived as follows. First, consider how often an undiscriminating female of genotype P_0 mates with T_i-males. In this case mating is random, and the numbers or frequencies of matings, n_i, adhere to a simple mass action law. They compute to

$$n_1(P_0) = y_1^{++} \tag{5.4a}$$
$$n_2(P_0) = y_2^{++} = 1 - y_1^{++}, \tag{5.4b}$$

where y_i^{++} is the frequency of males with trait T_i after viability selection. Note that the total number of matings of a P_0-female is one, $n(P_0) = \sum_i n_1(P_0) = 1$. To calculate the number of matings of choosy females, consider a P_1-female that encounters T_1-males with probability y_1^{++}, and T_2-males with probability y_2^{++}. In the former case, it accepts courtship and reproduction follows. In the latter case, it rejects courtship with probability p_1 which then entails a new mating round with probability q, but it proceeds to mate with the T_2-male with probability $1 - p_1$. That is to say, future mating rounds are discounted at a rate $y_2^{++}\, p_1\, q$. For the full numbers of matings, we track the chain of mating rounds and apply the summation formula for geometric series:

$$\begin{aligned} n_1(P_1) &= y_1^{++} + y_2^{++}\, p_1\, q \cdot y_1^{++} + \left(y_2^{++}\, p_1\, q\right)^2 \cdot y_1^{++} + \dots \\ &= y_1^{++} \left[1 + p_1\, q \left(1 - y_1^{++}\right) + p_1^2\, q^2 \left(1 - y_1^{++}\right)^2 + \dots\right] \\ &= \frac{y_1^{++}}{1 - p_1\, q \left(1 - y_1^{++}\right)}, \end{aligned} \tag{5.5a}$$

$$\begin{aligned} n_2(P_1) &= y_2^{++}\,(1 - p_1) + y_2^{++}\, p_1\, q \cdot y_2^{++}\,(1 - p_1) + \dots \\ &= (1 - p_1)\left(1 - y_1^{++}\right)\left[1 + p_1\, q \left(1 - y_1^{++}\right) + p_1^2\, q^2 \left(1 - y_1^{++}\right)^2 + \dots\right] \\ &= \frac{(1 - p_1)\left(1 - y_1^{++}\right)}{1 - p_1\, q \left(1 - y_1^{++}\right)}. \end{aligned} \tag{5.5b}$$

For P_2-females, a likewise calculation can be performed. Now, if we heed that the mating frequencies depend on the trait frequencies and thus have to be calculated separately for each population, we can finally merge them into the desired weight of sexual selection as follows:

$$S_{i,j,k} = \begin{cases} 1 & \text{if } j = 0 \\ \dfrac{1}{1 - p_j\, q \left(1 - y_{i,j}^{++}\right)} & \text{if } j = k \\ \dfrac{1 - p_j}{1 - p_j\, q \left(1 - y_{i,j}^{++}\right)} & \text{if } j > 0 \ \wedge \ j \neq k. \end{cases} \tag{5.6}$$

Here $y_{i,j}^{++}$ denotes the frequency of trait T_j after viability selection in population i, and $S_{i,j,k}$ yields the relative number of matings between P_j-females and T_k-males in that population. A comparison with Eq. (5.3) reveals the similarity of the two formulations of the sexual selection weight. Indeed, the similarity can be proven to be an equality for the case that $q = 1$. To see this, we first calculate the total number of matings a choosy P_j-female has:

$$n(P_j) = \sum_k n_k(P_j) = \frac{1 - p_j \left(1 - y_{i,j}^{++}\right)}{1 - p_j\, q \left(1 - y_{i,j}^{++}\right)}. \tag{5.7}$$

For a never-ending series of mating rounds, $q = 1$, numerator and denominator become equal, and it holds that $n(P_j) = 1$, so that a discriminating female has the same number of matings as a wildtype female. In that case, it is possible to map the rejection probability onto the mating preference strength and vice versa:

$$a_j = \frac{p_j}{1 - p_j}, \tag{5.8a}$$

$$p_j = \frac{a_j}{1 + a_j}, \tag{5.8b}$$

for $q = 1$. Mathematically put, the rejection probability is a transformation of the mating preference strength from the domain $[0, \infty]$ on the domain $[0, 1]$ which we consider more practical. Also note that if modeling of female mating preference is based on a preference strength parameter, no explicit assumptions are made as to how population-wide mating frequencies come about. It is as if females scan or sample the population for trait frequencies so as to know how to distribute their matings. In contrast, if female mating preference is modeled by a rejection probability parameter, population-wide mating frequencies follow directly from single female-male encounters and do not require any comprehensive population knowledge. In this study, we will nevertheless use the terms *mating preference strength* and *rejection probability* more or less synonymously, and use the variables a and p explicitly whenever a distinction is necessary.

If $q < 1$ then a choosy female has on average a smaller number of matings than a non-discriminating female. Indeed, we can use this difference in total mating numbers as a measure of the costs of a female harboring mating preference allele P_j in population i:

$$c(P_j) = 1 - n(P_j) = \frac{p_j\,(1 - q)\left(1 - y_{i,j}^{++}\right)}{1 - p_j\, q \left(1 - y_{i,j}^{++}\right)}. \tag{5.9}$$

In Fig. 5.2 the costs are plotted as a function of the frequency of the preferred trait for different values of the rejection and the transition probability. Costs monotonously decrease with the preferred trait's frequency. This is because unwanted males rarely need to be rejected due to the low probability of encountering them. Furthermore, increasing the transition probability as well as decreasing the rejection probability (which is equivalent to lowering the strength of the mating preference) reduce costs.

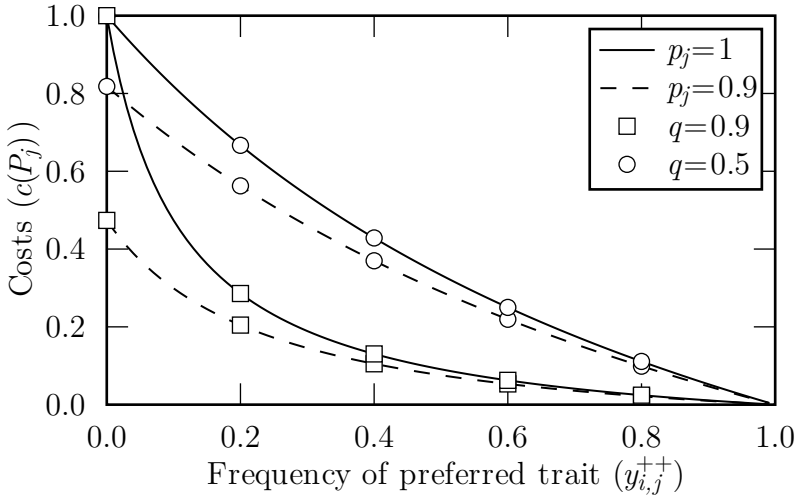


Figure 5.2: Costs of female mating preference. Females harboring the mating preference allele P_j in population i incur costs if the transition probability is less than one, $q < 1$. This is due to a reduced overall number of matings when they reject unwanted males with probability p_j. Costs are frequency dependent on the preferred trait, T_j, which females encounter at a rate $y_{i,j}^{++}$. Costs decline with increasing $y_{i,j}^{++}$, with decreasing rejection probability p_j, and with increasing transition probability q.

We have specified our modeling of costly female mating preference, but rather than setting up equations that describe the sexual selection step separately, it is more convenient to combine it with the reproduction step which we describe in the following.

Reproduction

The last step in the life cycle is the reproduction of new haploid offspring. This includes sexual selection, fecundity costs of infection, *Wolbachia* transmission, cytoplasmic incompatibility, and the inheritance of nuclear genes. In order to state the equations that model this last step, we first define the following weighting factors. *Female fecundity* is weighted according to

$$F_i = \begin{cases} 1-f & \text{if } i=1 \\ 1 & \text{if } i=2, \end{cases} \tag{5.10}$$

where i denotes the female cytotype, and where f are the fecundity costs caused by a *Wolbachia* infection.
Cytotype inheritance occurs from mother to offspring, and it is described by

$$V_{i,j} = \begin{cases} t & \text{if } i=1 \ \wedge \ j=1 \\ 1-t & \text{if } i=1 \ \wedge \ j=2 \\ 0 & \text{if } i=2 \ \wedge \ j=1 \\ 1 & \text{if } i=2 \ \wedge \ j=2, \end{cases} \tag{5.11}$$

where i and j are the cytotypes of female and offspring, respectively, and where t is the transmission rate of *Wolbachia*.
Cytoplasmic incompatibility reduces offspring numbers if uninfected eggs are fertilized by *Wolbachia*-modified sperm. The corresponding weight is

$$L_{i,j} = \begin{cases} 1 - l_{\mathrm{CI}} & \text{if } i{=}1 \ \wedge \ j{=}2 \\ 1 & \text{else.} \end{cases} \tag{5.12}$$

Here l_{CI} is called the level of cytoplasmic incompatibility, and it is defined as the proportion of zygotes that die if uninfected eggs are fertilized by sperm from infected males. Be aware that the indexes i and j are the cytotypes of father and offspring, respectively (and not of female and male, as one might expect). This is because infected females may also produce uninfected eggs due to imperfect *Wolbachia* transmission (see Eq. (5.11)) which are then compatible with sperm from infected males.
Nuclear inheritance of alleles at the trait and mating preference loci is accounted for by the weight

$$I_{i,j,k,l,m,n} = \begin{cases} 1 & \text{if } m{=}i{=}k \ \wedge \ n{=}j{=}l \\ \frac{1}{2} & \text{if } \big(m{=}i{=}k \ \wedge \ (n{=}j{\neq}l \ \vee \ n{=}l{\neq}j)\big) \\ & \qquad \vee \ \big((m{=}i{\neq}k \ \vee \ m{=}k{\neq}i) \ \wedge \ n{=}j{=}l\big) \\ \frac{1}{2}(1-r) & \text{if } \big(m{=}i \neq k \ \wedge \ n{=}j \neq l\big) \\ & \qquad \vee \ \big(m{=}k \neq i \ \wedge \ n{=}l \neq j\big) \\ \frac{1}{2}r & \text{if } \big(m{=}i \neq k \ \wedge \ n{=}j \neq l\big) \\ & \qquad \vee \ \big(m{=}k \neq i \ \wedge \ n{=}l \neq j\big) \\ 0 & \text{else,} \end{cases} \tag{5.13}$$

where the index pairs (i,j), (k,l), and (m,n) denote the alleles at the T and P loci of female, male, and offspring, respectively. In other words, the female genotype is (T_iP_j), the male genotype is (T_kP_l), and the offspring's genotype is (T_mP_n). The parameter r ($0 \leq r \leq \frac{1}{2}$) is the recombination rate and serves as a measure for the linkage between the two loci. If $r = 0$ then the loci are on the same chromosome and immediately adjacent to one another, and offspring always inherits either the mother's or the father's allele combination. At the other end of the r range, at $r = \frac{1}{2}$, alleles segregate independently because they are on different chromosomes or at opposite ends of the same chromosome.

From the frequencies of the nucleocytotypes after viability selection, $x^{++}_{i,j,k,l}$, we get the frequencies in the next generation by first summing over all possible matings, weighted by the factors $S_{i,j,k}$, F_i, $L_{i,j}$, $V_{i,j}$, and $I_{i,j,k}$, and then deviding by the overall number of offspring in the population, W_i:

$$x'_{i,j,k,l} = \frac{1}{W_i} \sum_{\alpha,\delta} \sum_{\beta,\varepsilon} \sum_{\gamma,\zeta} x^{++}_{i,\alpha,\beta,\gamma}(\female)\, x^{++}_{i,\delta,\varepsilon,\zeta}(\male)\, S_{i,\beta,\delta}\, F_\gamma\, L_{\zeta,l}\, V_{\gamma,l}\, I_{\alpha,\beta,\delta,\varepsilon,j,k}, \tag{5.14a}$$

$$W_i = \sum_{j,k,l} \sum_{\alpha,\delta} \sum_{\beta,\varepsilon} \sum_{\gamma,\zeta} x^{++}_{i,\alpha,\beta,\gamma}(\female)\, x^{++}_{i,\delta,\varepsilon,\zeta}(\male)\, S_{i,\beta,\delta}\, F_\gamma\, L_{\zeta,l}\, V_{\gamma,l}\, I_{\alpha,\beta,\delta,\varepsilon,j,k}. \tag{5.14b}$$

Equations (5.14) give the frequencies of offspring of nucleocytotype $T_jP_kC_l$ that is produced from matings between females of nucleocytotype $T_\alpha P_\beta C_\gamma$ and males of nucleocytotype $T_\delta P_\varepsilon C_\zeta$ in population i. We have marked the frequencies of females (♀) and males (♂) for clarity.

In summary of our mathematical formalization, the different life cycle steps of the host organisms are modeled by Equations (5.1) (migration), (5.2) (viability selection), and (5.14) (reproduction including sexual selection). Together these equations allow the calculation of the next generation nucleocytotype frequencies $x'_{i,j,k,l}$ from the current frequencies $x_{i,j,k,l}$. This system of equations, however, is far too complex to be handled analytically. Therefore, we conducted numerical simulations to explore the influence of parameter changes on the system dynamics. In particular, we were interested in the conditions that are necessary for one or both of the mating preference alleles P_1 and P_2 to spread, and under what circumstances this results in divergence at the P locus and thus reinforcement of a prezygotic isolation barrier.

5.2.3 Numerical simulations

Spread of the P_1 allele was considered to have occurred if its equilibrium frequency was above the introductory frequency of 0.001. An equilibrium was considered to be reached if in no populations any nucleocytotype frequency changed by more than 10^{-7} between subsequent generations, and if additionally overall change in frequencies (i.e., summed over all nucleocytotypes and populations) was declining. Parameter thresholds were in general determined with three-digit precision.

5.3 Results

We conducted extensive computer simulations of our mathematical model to investigate under which circumstances postzygotic isolation (caused by unidirectional CI) selects for premating isolation due to the evolution of female mating preferences.

We start the results section with a scenario in which an uninfected island population receives migration from a *Wolbachia* infected mainland (see Fig. 3.1b). Next, we analyze the processes that occur in two-way migration scenarios with a stable infection polymorphism (see Fig. 3.1c). For all scenarios, we determine equilibrium frequencies of the system after introduction of the preference allele P_1 in the uninfected population, examine how model parameters affect the system's behavior, and retrieve parameter thresholds that indicate qualitative changes. In some cases, we derive analytical heuristics which allow an intuitive understanding of the model. Finally, we quantify the impact of reinforcement on gene flow between populations following the numerical method described in Section 4.2.5.

Generally, the migration rate is chosen to be below the critical migration rate (compare Chapter 3), and selection on the trait locus is weak but sufficiently strong to maintain divergence between the populations in the face of migration. In all simulations performed, a minimal level of CI was necessary for P_1 to spread. In other words, in a "null model" without *Wolbachia*-induced CI, enhancing the production of locally better adapted offspring is not sufficient for P_1 to spread. Similarly, we found that female preference for T_2 males never evolves, i.e., neither in the scenario

with an uninfected mainland and an infected island nor in the two-way migration scenario does P_2 spread under any parameter constellation. Therefore, we will only consider alleles P_0 and P_1 in the remainder of the results section.

5.3.1 Uninfected island

In a scenario with an uninfected island population receiving migration from a *Wolbachia*-infected mainland, the mutant allele at the locus for female mating preference, P_1, spreads under a broad range of parameter values. Qualitatively similar results were presented by Servedio (2000) for a mainland-island model of reinforcement where postzygotic isolation is caused by nuclear epistatic interactions.

Equilibrium frequencies

After introduction in the mostly uninfected island population at a low frequency of 0.1%, the new mating preference allele may reach a stable equilibrium frequency which critically depends on the parameters of the system. For a certain set of parameter values ($f = 0.1$, $m = 0.0025$, $p_1 = 0.9$, $q = 1$, $r = 0.5$, $s = 0.1$, $t = 1$), Fig. 5.3a shows that P_1 goes to extinction if the CI level is below $l_{\mathrm{CI}} \approx 0.381$ but spreads if CI is above this threshold, reaching higher frequencies with increasing levels of CI. Note that for this parameter set, the level of CI that is necessary for *Wolbachia* to be stably maintained in the mainland population, is $l_{\mathrm{CI}} = 0.1$ (compare Eq. (3.4)). That is to say there exists a CI level range, $0.1 < l_{\mathrm{CI}} < 0.381$, where the infection polymorphism is stable but the mating preference allele can not spread.

The minimal CI level for P_1 to spread has its roots in two opposing forces. CI creates a selective advantage for P_1 but in order to spread this advantage must be sufficient to offset the gene flow of wildtype P_0 from the mainland. Note that choosiness itself creates a double advantage for P_1 females. First, mating with the locally better adapted T_1 males results in locally better adapted offspring. Second, matings with T_1 males result less often in CI because T_1 males are less likely to be infected with *Wolbachia*. As stated above, the former advantage is not sufficient to let P_1 spread, but the latter can be a huge advantage because P_1 females are mostly not infected with *Wolbachia*. From the female's perspective, the advantage of being choosy is therefore not only to find a male that fits the environment but predominantly to find a male that is compatible with one's own cytotype. Because the selected allele and cytotype are in association, female choice of the selected allele enhances chances of mating with a compatible cytotype.

The equilibrium frequency of the mating preference allele depends not only on the CI level but on all other parameters of the model. In fact, the pattern described for the dependence on the CI level can be found for most other parameters, as is illustrated in Fig. 5.4. The strength of the mating preference generated by the P_1 allele, p_1, resembles the CI level the most in its effect. The transmission rate of the *Wolbachia* infection, the transition probability of advancing to another mating round, and the strength of selection at the trait locus also show a threshold above which mating preference is established, and below which the wildtype allele P_0 prevails exclusively.

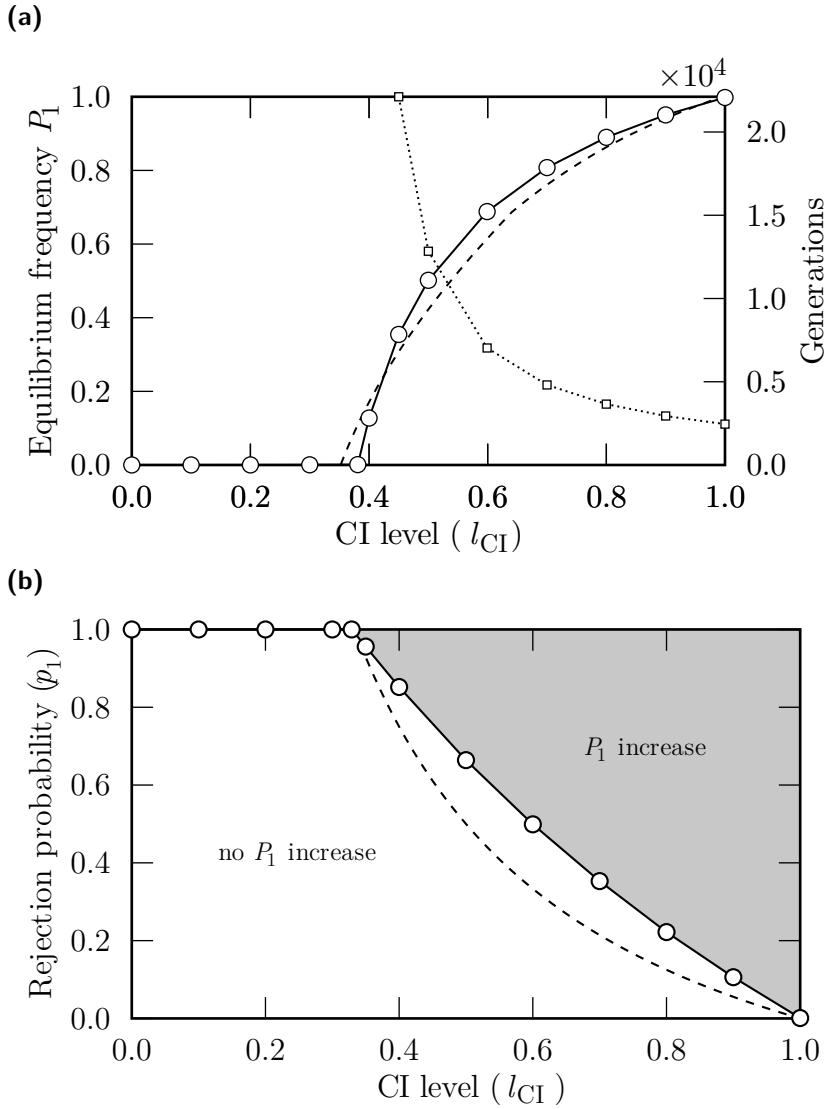


Figure 5.3: Equilibrium frequencies, rates, and thresholds of spread of the female mating preference allele. In both subgraphs, open symbols indicate numerically calculated values, and the dashed line gives the approximation according to Eq. (5.25). (a) Equilibrium frequencies (solid line with circles, left y–axis) of P_1 and number of generations (dotted line with squares, right y–axis) it takes the allele to propagate from 10% to 90% of the equilibrium frequency as a function of the level of CI. Spread occurs only if is strong enough, and it does so all the faster, the stronger CI becomes. (b) The parameter plane spanned by CI level and rejection probability of the P_1 allele consists of two regions. In the gray shaded region, P_1 spreads whereas it goes extinct in the white region. If $l_{\mathrm{CI}} > \frac{1}{3}$, then P_1 spreads on the island if the rejection probability is above a threshold. Parameters: (a) $f = 0.1$, $m = 0.0025$, $p_1 = 0.9$, $q = 1$, $r = 0.5$, $s = 0.1$, $t = 1$; (b) $f = 0.1$, $m = 0.001$, $q = 1$, $r = 0.5$, $s = 0.1$, $t = 1$.

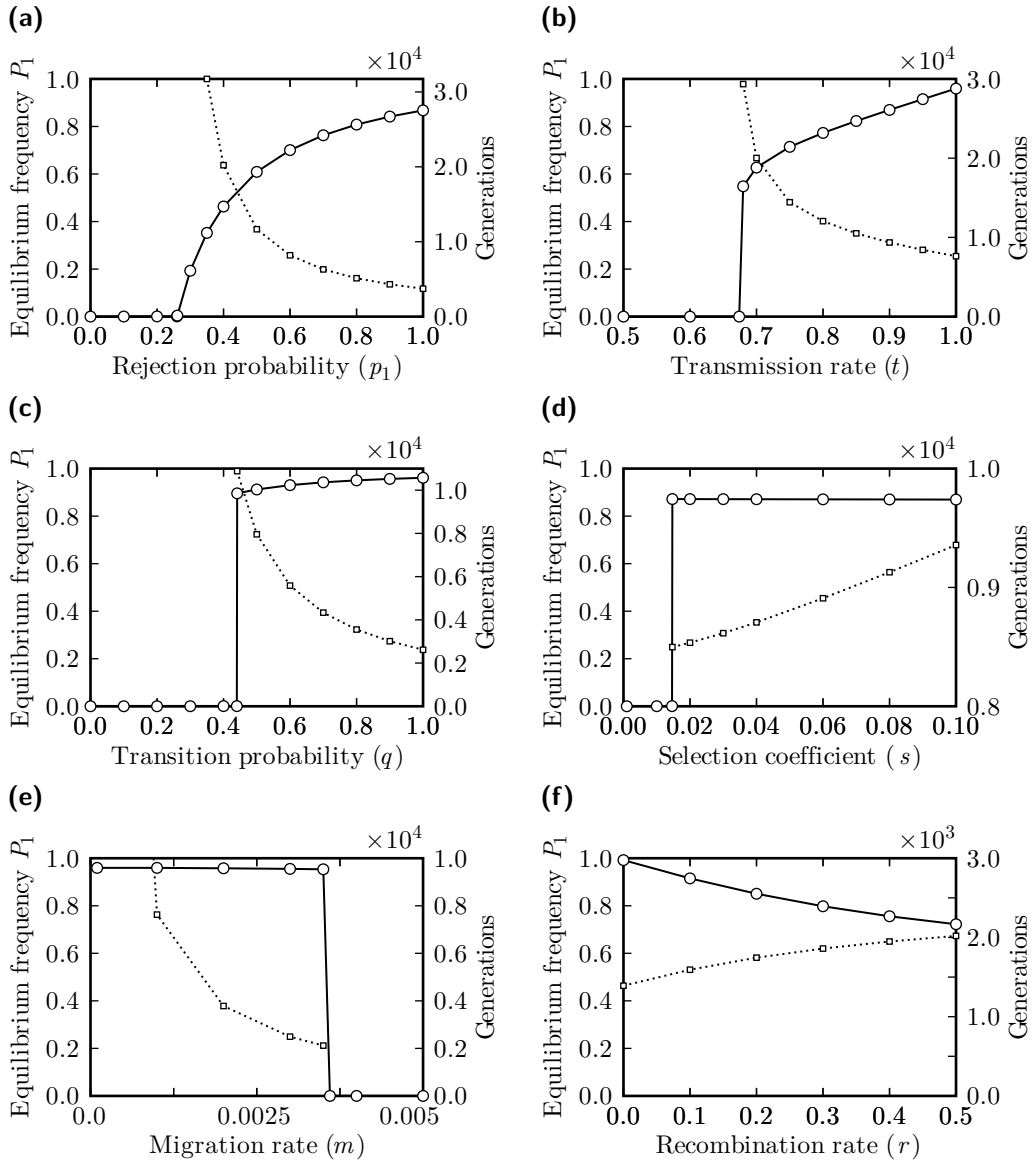


Figure 5.4: Equilibrium frequencies and rates of spread. Equilibrium frequencies (solid lines with open circles, left y–axis) of mating preference allele P_1 and number of generations it takes the allele to propagate from 10% to 90% of its equilibrium frequency (dotted line with squares, right y–axis) after introduction on the island with the low frequency of 0.1%. All values are numerically calculated and plotted against (a) the rejection probability, (b) the transmission rate, (c) the transition probability, (d) the selection coefficient, (e) the migration rate, and (f) the recombination rate. Parameters: (a) $m = 0.0025$, $q = 0.9$, $t = 0.9$; (b) $m = 0.001$, $q = 0.9$; (c) $m = 0.0025$, $t = 1$; (d) $m = 0.001$, $q = 0.9$, $t = 0.9$; (e) $q = 0.9$, $t = 0.9$; (f) $m = 0.01$, $q = 0.65$, $t = 0.9$; all subgraphs: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $p_1 = 1$, $r = 0.5$, $s = 0.1$ (if not on the x–axis).

Rates of spread

An important aspect of the possible evolution of premating isolation is the rate at which it occurs. If the time it takes to establish a new allele for female mating preference in our deterministic models is very long, it is likely that in real-world systems, other selective forces or genetic drift will prevent the spread of mating preference. Note that the total number of generations to the equilibrium state is not a good measure because the system dynamic is arbitrarily slow at the start and at the end of the spread, depending on the introduction frequency of the allele and the precision by which the equilibrium is considered to be reached (compare Section 5.2.3). In Figures 5.3a and 5.4, we have attempted to illustrate the rate of spread by plotting the number of host generations it takes for the mutant allele P_1 to propagate from 10% of its equilibrium frequency to 90%. By limiting the generation count in this manner, we neglect the arbitrarily long tails at both ends of the trajectory.

It is evident from the plots that nowhere in the parameter space does the female mating preference spread considerably fast. Typical times to bridge the $10-90\%$ range are in the order of magnitude of thousands to tens of thousands of host generations. To put these numbers in relation to direct selection, this is roughly equivalent to the rate of spread of an allele that is favored by local selection by a factor of 1.0001 to 1.001 in the face of migration by a rate of 10^{-5} to 10^{-4}.

The system dynamics become very slow if CI level or rejection probability are just above the critical parameter thresholds (see Figures 5.3a and 5.4a, respectively). With transition probability and transmission rate, we observed a similar but less pronounced slow-down close to critical values (see Figures 5.4b and 5.4c, respectively). In all of these four cases, female mating preference spreads most rapidly for large parameter values. In contrast, a decrease of the strength of selection at the trait locus or of the recombination rate between trait and preference loci both produce a faster spread of the P_1 allele (see Figures 5.4d and 5.4f, respectively). Note however that these effects are comparably small. Finally, the migration rate accelerates the P_1 dynamic but only if it stays below the critical migration rate (see Fig. 5.4e).

In conclusion, in order for female mating preference to spread in the island population at appreciable rates, the preference should be strong and associated with little costs. Furthermore, CI level and transmission rate of *Wolbachia* should be close to one, and migration should be as strong as possible without destroying the infection polymorphism. Strength of trait selection and the rate of recombination can be neglected in this context.

Bifurcation analysis

The plots of P_1 equilibrium frequencies as functions of model parameters as in Fig. 5.3a and Figures 5.4 can be interpreted as bifurcation diagrams of a one-dimensional discrete time system. The following classification of bifurcations can be made. With respect to l_{CI} and p_1, supercritical pitchfork bifurcations occur at the threshold values, with a new branch of stable fixpoints forking from the fixpoint $x^*_{P_1}=0$ which becomes unstable (compare Fig. 5.5a). In the context of our biologically motivated model, only the upper branch is meaningful and determines

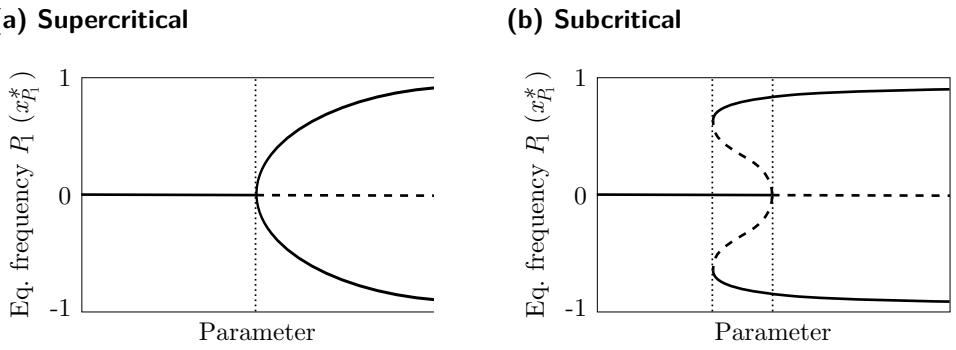


Figure 5.5: Pitchfork bifurcations. (a) The mainland-island system undergoes a supercritical pitchfork bifurcation for parameters l_{CI} and p_1. (b) For parameters q and s, subcritical pitchfork bifurcations occur. Stable fixpoint branches are given by solid lines, unstable fixpoints by dashed lines. Only the upper branches are biologically meaningful because the frequency of P_1 must lie within the range $[0, 1]$. The dotted lines mark the threshold (or critical) parameter values.

the equilibrium frequency of the mating preference allele P_1. The bifurcations occurring with respect to parameters q and s are of the subcritical pitchfork type (compare Fig. 5.5b). Here, a new branch of unstable fixpoints is created at the threshold values which subsequently undergoes a saddle-node bifurcation (at a lower parameter threshold), thus changing the stability of the branch. Note that this kind of bifurcation implies a hysteresis effect when tracking the upper stable branch backwards (i.e., decreasing the parameter).

Because our model scenario includes a low starting frequency of the mutant allele so that we always start close to the fixpoint $x^*_{P_1} = 0$, only the larger threshold value is captured, and the equilibrium frequency abruptly jumps to the upper stable branch (compare Figures 5.4c, 5.4d). We verified the existence of the branches that lie between the two threshold values (i.e., between the dotted lines in Fig. 5.5b) by numerically simulating the system with the P_1 allele fixed in the beginning (results not shown). For the set of parameter values displayed in Fig. 5.4c, the threshold at which the saddle-node bifurcation occurs with respect to the transition probability is at $q \approx 0.24$ which is considerably lower than the pitchfork bifurcation threshold at $q \approx 0.44$. It is interesting to note, that for the set of parameter values in Fig. 5.4d, the saddle-node bifurcation does not occur for biologically relevant values of the selection coefficient at all but only at $s < 0$. This implies that once mating preference has evolved, selection at the trait locus might no longer be necessary to maintain genetic divergence.

The parameters that differ from the described bifurcation patterns are the rate of migration, the transmission rate of *Wolbachia*, and the rate of recombination between the nuclear loci T and P (see Figures 5.4e, 5.4b, and 5.4f, respectively). With respect to the migration rate, m, the infection polymorphism breaks down if a threshold value is exceeded. This is the critical migration rate we have analyzed in detail in Chapter 3 (see Eq. (3.10)), and it is the saddle-node bifurcation of the *Wolbachia* frequency that is reflected in the branching of $x^*_{P_1}$ here. This is a further affirmation of the rule that spread of P_1 is only possible if the infection

polymorphism is stable. The threshold value of the transmission rate, t, is also determined by the stability of the infection on the mainland. In this sense, it is equivalent to the critical CI level we have calculated in Eq. (3.4a). Finally, for large parts of the relevant parameter space, changing the recombination rate does not result in any bifurcation (compare Fig. 5.4f). The only cases we have found where (saddle-node) bifurcations occur when r exceeds a threshold are if other parameters are such that fixpoint branches exist that are not reachable from low starting frequencies (as described above for parameters q and s). Note however that the stability of the $x^*_{P_1} = 0$ branch is not affected by these bifurcations so that the spread of a new mutant at the preference locus can never be made possible only by close physical linkage to the trait locus.

Thresholds for premating isolation

We screened the parameter space more generally to detect thresholds for divergence at the preference locus. In Fig. 5.3b, the parameter plane spanned by the CI level and the rejection probability of the mutant allele is shown. It demonstrates that a spread of P_1 is not possible for low CI levels no matter how strong the mating preference is. However, if the level of CI is above a threshold, P_1 spreads on the island if mating preference is strong enough. The minimal rejection probability necessary for P_1 to spread decreases with increasing CI levels. Fig. 5.6 illustrates that similar thresholds exist for (a) the transmission fidelity of *Wolbachia*, (b) the probability of transition into a new mating round, (c) the coefficient of selection acting on the trait locus, and (d) the rate of recombination between trait and preference loci. The following conclusions can be drawn from the thresholds. (a) P_1 can only spread if *Wolbachia* transmission is above a threshold rate because otherwise the infection is lost in the mainland population. However, the greater the transmission fidelity is, the greater becomes the risk of CI matings; therefore, the more advantageous it becomes to be choosy so that less strong mating preferences suffice for a spread of P_1. (b) It is evident that increasing costs of mating preference (by decreasing the transition probability, q) requires P_1 to generate stronger preference in order to spread. If the transition probability drops below a threshold then the P_1 allele can not spread any longer because mating preference has become too costly. (c) Selection on the trait locus must be strong enough to allow the spread of the mating preference mutant. (d) Reduced recombination between trait and mating preference loci can significantly enhance equilibrium frequencies of P_1 and lower the minimal rejection probability that is necessary for the mutant allele to spread, even if costs are considerable. This is because the association of trait, preference and infection status to form the nucleocytotype T_1P_1U which is favored the most by natural and sexual selection on the island is facilitated by the physical linkage of preference and trait loci on a host chromosome.

Analytical considerations

As discussed above, there are two opposing forces acting on the preference mutant allele P_1 in the island population. On the one side, P_1 individuals have a selective advantage over P_2 individuals because they are less often involved in incompatibility matings. On the other side, there is permanent gene flow of the P_2 allele from the

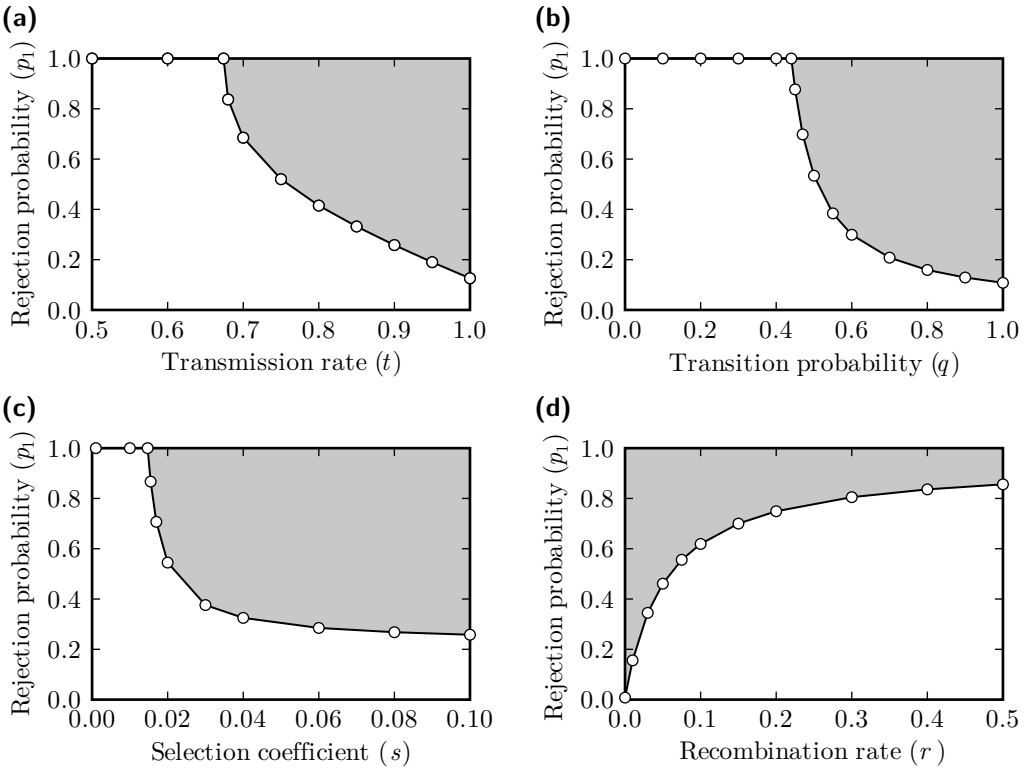


Figure 5.6: Thresholds for the spread of the mating preference mutant. The parameter planes spanned by the rejection probability of the P_1 allele on the y–axis and (a) the transmission rate, (b) the transition probability, (c) the selection coefficient, or (d) the recombination rate on the x–axis all consist of two regions. In gray shaded regions, P_1 spreads whereas it goes extinct in white regions. Open circles indicate values calculated by numerical simulations. Parameters: (a) $m = 0.001$, $q = 0.9$, $r = 0.5$, $s = 0.1$; (b) $m = 0.0025$, $r = 0.5$, $s = 0.1$, $t = 1$; (c) $m = 0.001$, $q = 0.9$, $r = 0.5$, $t = 0.9$; (d) $m = 0.01$, $q = 0.65$, $s = 0.1$, $t = 0.9$; all subgraphs: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$.

mainland. Here, we formalize the verbal reasoning and derive heuristic formulas for the calculation of P_1 allele frequencies.

We use the function $F = F(x, m, s)$, defined in Eq. (4.30), to describe selection and migration acting on the preference locus. In order to apply the function F to the P_1 dynamics we will define an "effective selection coefficient" s_{eff} that describes the selective advantage of P_1 over the wildtype allele P_0. In addition, we take into account that gene flow is appropriately described by the effective migration rate m_{eff} due to cytoplasmic incompatibility. As we will show, the following equations are good approximations for the dynamics and equilibrium frequencies of the P_1 allele,

$$x'_{P_1} \approx F(x_{P_1}, m_{\mathrm{eff}}, s_{\mathrm{eff}})\,, \tag{5.15a}$$

$$x^*_{P_1} \approx G(m_{\mathrm{eff}}, s_{\mathrm{eff}}) \approx 1 - \frac{m_{\mathrm{eff}}}{(1 - m_{\mathrm{eff}})\, s_{\mathrm{eff}}}, \tag{5.15b}$$

Here, x_{P_1} and x'_{P_1} denote the frequencies of the P_1 allele on the island in subsequent generations, and $x^*_{P_1}$ its equilibrium frequency. $G(m, s)$ is the solution function of

the allele dynamics as given by Eq. (4.31).

In order to define an effective selection coefficient for a rare x_{P_1} allele in a x_{P_0} population we determine the average fitness of the respective alleles. First, we consider the case that x_{P_1} females mate exclusively with T_1 males, i.e., $p_1 = 1$ which is equivalent to $a_1 = \infty$. Furthermore, we assume that *Wolbachia* transmission is perfect, $t = 1$, and that there are no costs of female mating preference due to a transition probability of one, $q = 1$. Under these assumptions, there is a strong linkage disequilibrium between P_1 and T_1. Let x_W denote the frequency of *Wolbachia* on the island. Then the average fitness of P_0 is $1 - l_{\mathrm{CI}}\, x_W$. The fitness of P_1 individuals is $1 - l_{\mathrm{CI}} \frac{x_{T_{1,W}}}{x_{T_1}}$ where x_{T_1} denotes the frequency of T_1 on the island and $x_{T_{1,W}}$ the frequency of infected T_1 individuals. As the effective selection coefficient we take the difference of both,

$$s_{\mathrm{eff}}(P_1) \approx l_{\mathrm{CI}} \left(x_W - \frac{x_{T_{1,W}}}{x_{T_1}} \right). \tag{5.16}$$

In order to get a useful approximation of $s_{\mathrm{eff}}(P_1)$, we assume that $x_{T_1} \approx 1$. This is justified for small migration rates. Denoting the frequency of infected T_2 individuals with $x_{T_{2,W}}$, we get

$$s_{\mathrm{eff}}(P_1) \approx l_{\mathrm{CI}}\, x_{T_{2,W}}. \tag{5.17}$$

We can approximate $x_{T_{2,W}}$ by considering the matriline of rare migrants from the mainland where the nucleocytotype T_2W is fixed. Immediate immigrants constitute a fraction m of the island population. The infected part of F_1 offspring of these migrants is $m\,(1-m)\,(1-f)/2$, and subsequent generations can be computed accordingly. Summing over the unfolding geometric series and linearizing around $m = 0$ gives

$$x_{T_{2,W}} \approx m + m\,\frac{(1-f)\,(1-m)}{2} + m\left(\frac{(1-f)\,(1-m)}{2}\right)^2 + \cdots \tag{5.18a}$$

$$\approx \frac{2\,m}{1 + f + (1-f)\,m} \tag{5.18b}$$

$$\approx \frac{2\,m}{1+f} \tag{5.18c}$$

Substitution into Eq. (5.17) yields the following useful approximation for the effective selection coefficient which depends only on the parameters of the system,

$$s_{\mathrm{eff}}(P_1) \approx \frac{2\,l_{\mathrm{CI}}\,m}{1+f}. \tag{5.19}$$

In a next step, we relax the above assumptions and consider the case where both rejection and transition probability are allowed to be smaller than one: $p_1, q < 1$. Note that the selective advantage of P_1 over P_0 is reduced by both relaxations. This is because P_1 females are more likely to be involved in incompatibility matings if p_1 is low, and because their overall number of matings is reduced if $q < 1$. In general, the fitness of a female can be described by

$$\text{fitness} = (1 - l_{\mathrm{CI}}\, x_{\mathrm{CI}}) \cdot n_{\mathrm{total}} \tag{5.20a}$$

$$= n_{\text{total}} - l_{\text{CI}}\, n_{\text{CI}},$$

where x_{CI} denotes the fraction of CI matings (of all matings of the focal female), and where n_{CI} and n_{total} are the absolute numbers of CI matings and total matings of that female, respectively. Note that for a wildtype P_0 female, it holds that $n_{\text{total}} = 1$ and $x_{\text{CI}} = x_W$ so that the above used formula $1 - l_{\text{CI}}\, x_W$ ensues. Drawing on Equations (5.5), we can compute the mating numbers for a rare P_1 allele:

$$n_{\text{total}}(P_1) = \frac{1 - p_1\,(1 - x_{T_1})}{1 - p_1\, q\,(1 - x_{T_1})}, \tag{5.21a}$$

$$\begin{aligned} n_{\text{CI}}(P_1) &= n_{T_{1,W}} + n_{T_{2,W}} \\ &= \frac{x_{T_{1,W}}}{1 - p_1\, q\,(1 - x_{T_1})} + \frac{(1 - p_1)\, x_{T_{2,W}}}{1 - p_1\, q\,(1 - x_{T_1})} \\ &= \frac{x_W - p_1\, x_{T_{2,W}}}{1 - p_1\, q\,(1 - x_{T_1})}. \end{aligned} \tag{5.21b}$$

Again, we set as the effective selection coefficient the fitness difference of the two mating preference alleles. This results in

$$s_{\text{eff}}(P_1) \approx \frac{l_{\text{CI}}\, p_1\, x_{T_{2,W}} - p_1\,(1 - x_{T_1})\,[1 - q\,(1 - l_{\text{CI}}\, x_W)]}{1 - p_1\, q\,(1 - x_{T_1})}. \tag{5.22}$$

Note that if $p_1 = 1$ and $q = 1$, Eq. (5.22) correctly reduces to Eq. (5.16).

Effective selection coefficient. Next, we use the approximations $x_{T_{2,W}} \approx \frac{2\,m}{1+f}$ (see Eq. (5.18)), $x_{T_1} \approx 1 - \frac{m}{s}\,\frac{1-l_{\text{CI}}}{1+f}$ as per Eq. (4.32b), and $x_W \approx \frac{m}{f}$ which is the *Wolbachia* frequency from Eq. (3.28b) approximated for weak migration by linearization around $m = 0$. This yields

$$s_{\text{eff}}(P_1) \approx \frac{[2\, l_{\text{CI}}\, s - (1 - l_{\text{CI}})\,(1 - q)]\, m\, p_1}{(1 + f)\, s}, \tag{5.23}$$

which only depends on the system parameters. Finally, to achieve approximations for P_1 equilibrium frequencies, we substitute this effective selection coefficient and the appropriate effective migration rate, $m_{\text{eff}} = \frac{1-l_{\text{CI}}}{1+f}\, m$ (see Eq. (4.17)), in Eq. (5.15b) and get

$$x^*_{P_1} \approx 1 - \frac{f\,(1 - l_{\text{CI}})\,[(1 - l_{\text{CI}})\, m\, p_1\, q - (1 + f)\, s]}{[1 + f - (1 - l_{\text{CI}})\, m]\,[(1 - l_{\text{CI}})\,(f\,(1 - q) + l_{\text{CI}}\, m\, q) - 2\, f\, l_{\text{CI}}\, s]\, p_1}, \tag{5.24}$$

For small migration rates, Eq. (5.24) greatly simplifies, and we get a practical approximation of the equilibrium frequency of the P_1 allele,

$$x^*_{P_1} \approx 1 - \frac{(1 - l_{\text{CI}})\, s}{[2\, l_{\text{CI}}\, s - (1 - l_{\text{CI}})\,(1 - q)]\, p_1}. \tag{5.25}$$

Threshold. A consequence of Eq. (5.25) is that P_1 can spread into the population only if the mating preference is strong enough,

$$p_1 > \frac{(1 - l_{\mathrm{CI}})\, s}{2\, l_{\mathrm{CI}}\, s - (1 - l_{\mathrm{CI}})\,(1 - q)}, \tag{5.26}$$

but goes to extinction if the rejection probability is below this threshold value. Note that if mating preference is not costly, $q = 1$, then Eq. (5.25) reduces to $x^*_{P_1} \approx 1 - \frac{1-l_{\mathrm{CI}}}{2\, l_{\mathrm{CI}}\, p_1}$, and the p_1 threshold is simply $\frac{1-l_{\mathrm{CI}}}{2\, l_{\mathrm{CI}}}$. Thereby, formulas (16) and (17) in Telschow et al. (2007) are correctly reproduced. The threshold p_1 probability decreases with increasing CI levels, with stronger selection on the preferred trait, and with increasing transition probabilities, i.e., with lower costs. We can compute the CI level for which complete discrimination, $p_1 = 1$, is just strong enough for the mutant allele to spread by letting the right hand side of Eq. (5.26) equal one and solving for l_{CI}. This yields $l_{\mathrm{CI}} = \frac{1-q+s}{1-q+3\, s}$. Thus, if mating preference does not incur costs on choosy females, reinforcement in principle is possible if the level of CI is larger than $\frac{1}{3}$. The dashed lines in Fig. 5.3 demonstrate that Equations (5.25) and (5.26) approximate the numerically determined values reasonably well. The approximations are especially good when mating preference is strong and CI levels are close to one.

5.3.2 Fisherian runaways

We now allow for migration to occur (symmetrically) in both directions and consider a two-population scenario with a stable infection polymorphism. We start with the case that female mating preference is not associated with costs, i.e., $q = 1$. In this case, female mating preference spreads even more easily than in the mainland-island scenario of the previous Section 5.3.1. But as readily as this spread results in premating isolation between an uninfected island and an infected mainland, as readily does it trigger a runaway process in the two-way migration scenario thus foreclosing premating isolation.

Two types of runaways

In Fig. 5.7, the equilibrium frequencies of trait and preference alleles T_1 and P_1 in both populations are plotted as functions of the strength of mating preference, p_1. If mating preference is negligibly weak (i.e., if the rejection probability is smaller than 1%), P_1 does not spread at all. However, if it stronger, runaway selection invariably occurs. Two types of runaway processes can be distinguished. First, if mating preference strength is below a threshold, $p_1 \lesssim 18\%$, the mutant allele P_1 becomes fixed in both populations, and the preferred trait, T_1, spreads to some extent in the infected population. Secondly, if $p_1 \gtrsim 18\%$ then T_1 becomes fixed in both populations, and P_1 spreads to the same equilibrium frequency in both populations. As a result, in neither case does divergence at the locus for female mating preference follow the spread of P_1: premating isolation does not evolve. Note that in both cases, the infection polymorphism is essentially unaffected.

Spread of female mating preference occurs because of indirect selection, and can be explained in the uninfected population by a variant of the "good gene" approach, and in the infected population by a "runaway process" (Kirkpatrick and Barton,

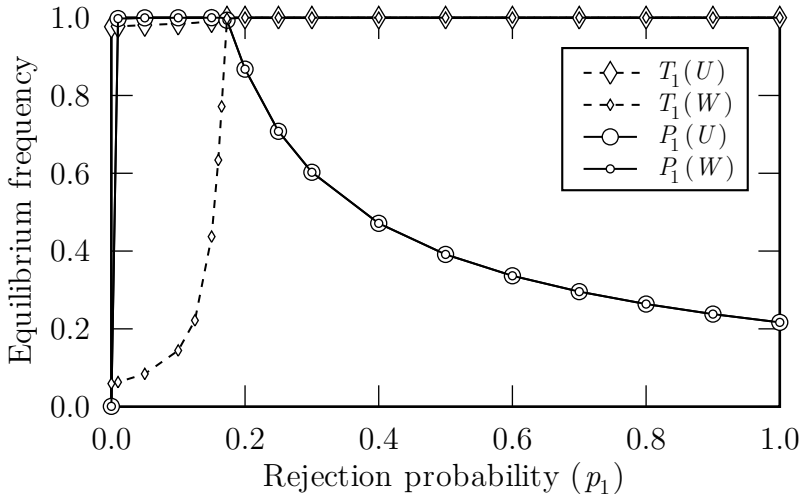


Figure 5.7: Fisherian runaways – Female mating preference without costs. In both the mainly uninfected population (U, large symbols) and the mainly infected population (W, small symbols), equilibrium frequencies of the mutant allele for female mating preference (P_1, circles) and of the preferred trait (T_1, diamonds) depend on the strength of the mating preference (p_1). There are three ranges of p_1 that result in qualitatively different system equilibria. If p_1 is below a certain yet negligibly small rejection probability (below 0.001) then P_1 goes to extinction. However, if p_1 is larger than this threshold then P_1 spreads and spawns a runaway process that comes in two flavors. If $p_1 \lesssim 0.18$ then P_1 spreads to fixation in both populations helping T_1 reach increasing equilibrium frequencies in the infected population with stronger preference. If $p_1 \gtrsim 0.18$ then T_1 spreads to fixation, and P_1 reaches the same frequency in both populations. In neither case does premating isolation between the populations evolve. Parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.01$, $q = 1$, $r = 0.5$, $s = 0.1$, $t = 0.9$.

1997). In the uninfected population, the trait of a male can be used by females as a cue for the infection status of the focal male because trait and cytotype are in linkage disequilibrium (more specifically, T_1 and $C_1 = U$ are positively linked). Thus, uninfected females that prefer T_1 males over T_2 males produce more viable offspring than uninfected wildtype females by reducing the risk of CI matings. In this sense, T_1 acts as a "good gene" in the uninfected population by indicating the viability of potential offspring. By contrast, in the infected population, no benefit arises from using the trait as an infection cue because females are infected themselves. However, because some P_1 females immigrate into the infected population, sexual selection in favor of T_1 males is created, in opposition to viability selection which favors T_2. Therefore, preferentially mating with T_1 males yields less viable but more "attractive" offspring, and a runaway process ensues.

The two types of runaway processes that can be observed in Fig. 5.7 stem from the fact that the ability to discriminate between males with different traits is of no avail if only males of the same trait remain. Consider the following line of arguments. At rejection probabilities below the threshold, P_1 is fixed in both populations, and a balance between sexual and viability selection (and migration) determines the trait frequencies in the infected population. With increasing preference strength of the P_1 allele, higher frequencies of the preferred trait in the unfavorable environment

are attained. At some point, sexual selection is so strong that viability selection is fully offset, and the preferred trait becomes fixed in both populations. In turn, the ability to discriminate between males is rendered pointless, P_1 becomes neutral and its spread is halted. The stronger the mating preference is, the faster the dynamics become, the more rapid fixation of T_1 is brought about, and the sooner the spread of P_1 comes to a halt (i.e., at lower frequencies). It should be kept in mind that the infection polymorphism forms the foundation that the runaway processes are built upon. However, *Wolbachia*'s influence on the type of runaway process is small as we will show below.

Analytical considerations

It is possible to derive a heuristic formula for the threshold of mating preference strength that determines the type of runaway process. Similar to the previous Section 5.3.1, our approach is to use the function $F = F(x, m, s)$ to describe T_2 allele dynamics in the infected population, and to develop an effective selection coefficient that encompasses viability selection and sexual selection created by female mating preference:

$$x'_{T_2} \approx F(m, s_{\text{eff}}, x_{T_2})\,. \tag{5.27}$$

Here, x_{T_2} and x'_{T_2} are the frequencies of trait T_2 in the infected population in consecutive generations. The key point in order to derive an effective selection coefficient is to compare the numbers of matings that males acquire. Let us assume that P_1 is fixed in both populations (i.e., p_1 is below the threshold value we want to approximate), and let x_{T_1} and x_{T_2} denote the frequencies of the two traits in the infected population. Using the weight that describes sexual selection as per Eq. (5.6), we can compute the relative numbers of matings of a T_1 and a T_2 male:

$$n(T_1) = \frac{1}{1 - p_1\,(1 - x_{T_1})} = \frac{1}{1 - p_1 x_{T_2}}, \tag{5.28a}$$

$$n(T_2) = \frac{1 - p_1}{1 - p_1\,(1 - x_{T_1})} = \frac{1 - p_1}{1 - p_1 x_{T_2}}. \tag{5.28b}$$

An effective selection coefficient for the locally adaptive trait T_2 can then be set up as

$$\begin{aligned} s_{\text{eff}}(T_2) &\approx s + n(T_2) - n(T_1) \\ &\approx \frac{s - p_1\,(1 + s x_{T_2})}{1 - p_1 x_{T_2}}. \end{aligned} \tag{5.29}$$

Inserting this effective selection coefficient into Eq. (5.27) and solving for $x^*_{T_2} = x'_{T_2} = x_{T_2}$ yields

$$\begin{aligned} x^*_{T_2} \approx \frac{(1-m)\,(1+p_1)\,s - p_1}{2\,(1-m)\,p_1\,s} \;-\; \\ \frac{\sqrt{p_1^2 - 2\,[1 - m\,(3 - 2\,m)]\,(1-p_1)\,p_1\,s + (1-m)^2\,(1-p_1)^2\,s^2}}{2\,(1-m)\,p_1\,s}. \end{aligned} \tag{5.30}$$

It holds that $x^*_{T_1} = 1 - x^*_{T_2}$, so we can use Eq. (5.30) to approximate the rise of the T_1 frequency in the infected population when the rejection probability is increased (small diamond symbols in Fig. 5.7).

Critical strength of mating preference. We can estimate the threshold for the rejection probability of P_1 females that marks the fixation of the preferred trait by setting $x^*_{T_2} = 0$ and solving for p_1:

$$p_{1,\text{crit}} \approx s - \frac{m}{1-m}. \tag{5.31}$$

Because migration rates are typically small if the infection polymorphism is to be stable (compare Chapter 3), the second term on the right hand side of Formula (5.31) can be neglected, and we can finally produce the following simple "rule of thumb":

$$p_{1,\text{crit}} \approx s. \tag{5.32}$$

If we heed that p_1 by definition can only range from zero to one, we can deduce from Eq. (5.32) that the type of runaway process that results in fixation of the preferred trait can only arise if viability selection at the trait locus occurs with coefficients smaller than one. Otherwise, mating preference will spread to fixation. Numerical simulations support this assertion (data not shown).

The simplifying step from Eq. (5.31) to Eq. (5.32) also shows that the influence that *Wolbachia* has on the critical p_1 threshold is small as CI results in migration being effectively decreased (see Chapter 4). The stability of the infection polymorphism, however, remains fundamental to the runaway process itself because the mating preference allele P_1 would not spread in the first place without the benefit of avoiding CI matings, as we have shown in the previous Section 5.3.1.

5.3.3 Reinforcement

In our model, females only get another chance to secure a mating partner with probability q (see the model section on sexual selection on page 92). So far, we have assumed that $q = 1$ so that eventual mating was certain even for a very choosy female in a population with very few males with the preferred trait. We will now analyze the case that the transition probability is smaller than one, $q < 1$, so that rejecting potential mating partners will reduce the number of matings a female acquires. The transition probability can therefore be used as a proxy for the costs of being choosy. Importantly, because of the frequency dependence on the preferred trait, costs of P_1 females are considerably smaller in the uninfected than in the infected population.

In Fig. 5.8, the equilibrium frequencies of the trait allele T_1 and the mutant allele for female mating preference P_1 are plotted against the strength of the mating preference for $q = 0.9$. Note that all other parameter values are the same as in Fig. 5.7 (which showed the inevitability of Fisherian runaways when mating preferences are not costly because $q = 1$). If the mating preference strength is weak, P_1 does not spread in neither population. However, if it is above a threshold, $p_1 \gtrsim 0.22$, P_1 spreads in the uninfected population but not (or only to a much lesser extent) in the infected one. Premating isolation increases with the strength of the

mating preference. Thus, the simple assumption of comparably small costs readily yields reinforcement of a premating barrier due to the CI barrier.

Favorable conditions for reinforcement. Reinforcement occurs under a broad range of parameter conditions. Several effects can be observed in our simulations: (i) The strength of reinforcement (i.e., the difference in P_1 equilibrium frequency between the two populations) increases with the strength of the isolation due to CI, i.e., with higher levels of CI. (ii) As long as the infection polymorphism is stable, higher *Wolbachia* transmission fidelity and larger fecundity costs of infection both yield stronger reinforcement. (iii) Reduced recombination between the trait and preference loci favors reinforcement. In the limit case of no recombination, $r = 0$, P_1 can practically reach fixation in the uninfected population. Under all other parameter constellations, however, fixation is never attained. (iv) The strength of selection at the trait locus must be above a threshold. (v) The transition probability must lie within a certain range, i.e., costs must neither be too small nor too high. We will expand on the latter two points in the following.

From runaway to reinforcement. Two parameters prove to be critical in determining whether reinforcement or a runaway process take place, the coefficient of local selection on the trait, s, and the transition probability, q. In Fig. 5.9, equilibrium frequencies of T_1 and P_1 are plotted against these two parameters (for a set of otherwise fixed parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.01$, $p_1 = 0.9$, $r = 0.5$, $t = 0.9$). Subgraph 5.9a shows that local selection must be above a threshold for reinforce-

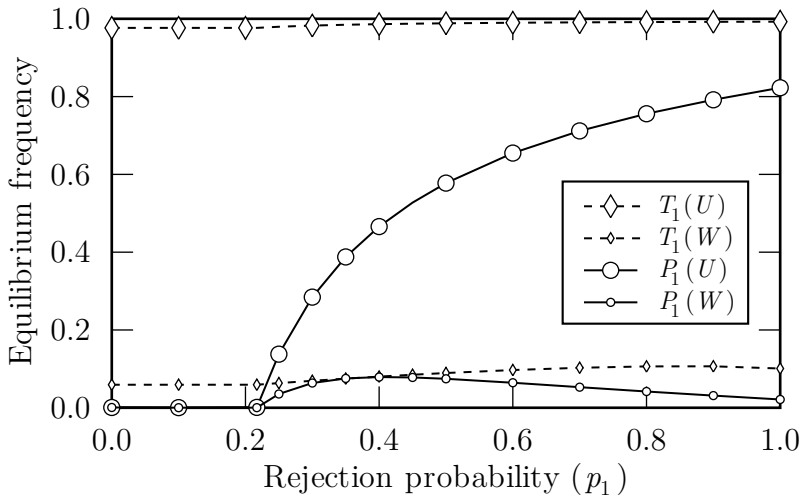


Figure 5.8: Reinforcement – Costly female mating preference. If female mating preference is costly for discriminating females, premating isolation between the mainly uninfected population (U, large symbols) and the mainly infected population (W, small symbols) readily evolves if the rejection probability is above a threshold, $p_1 \gtrsim 0.22$. Then, the mutant allele P_1 (circles) spreads to higher frequencies in the uninfected than in the infected population while divergence at the trait locus is maintained. The only difference to the scenario from Fig. 5.7 is that here, $q = 0.9$ instead of $q = 1$. Parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.01$, $q = 0.9$, $r = 0.5$, $s = 0.1$, $t = 0.9$.

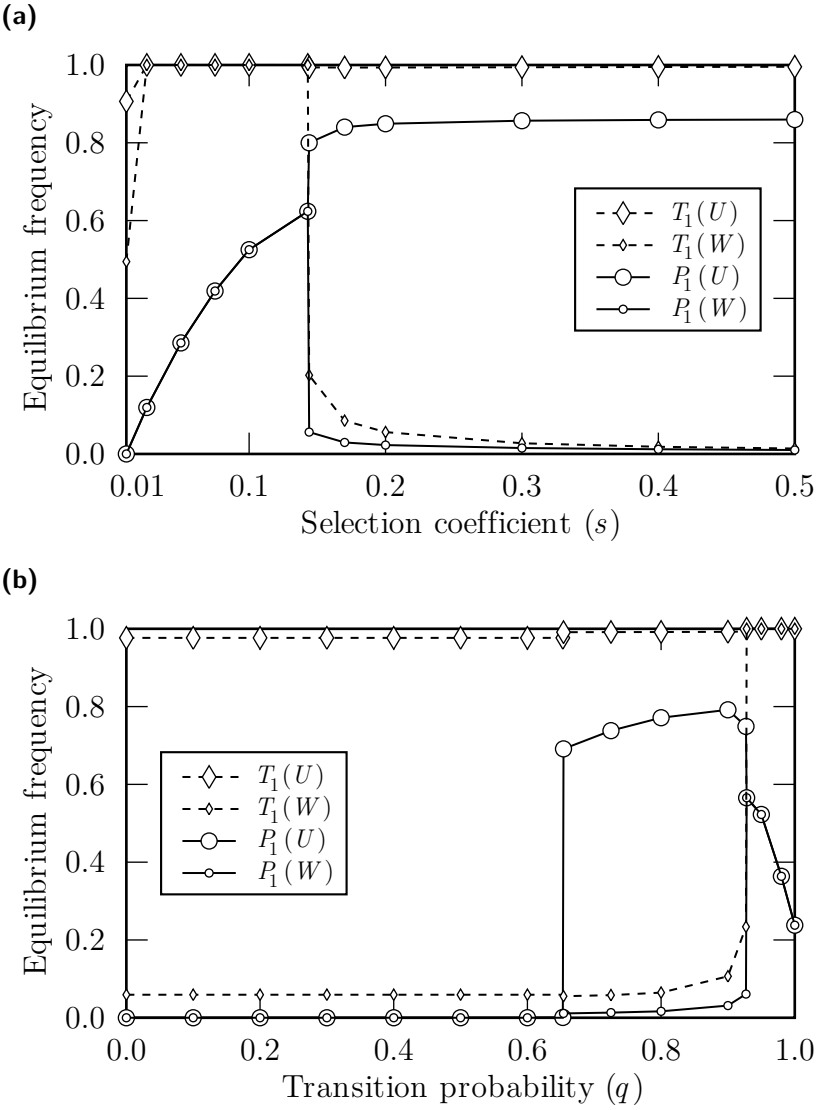


Figure 5.9: From runaway to reinforcement. (a) The spread of the allele for female mating preference, P_1 (circles), creates sexual selection favoring males carrying trait T_1 (diamonds) that are locally adaptive in the uninfected population (U, large symbols) but maladaptive in the infected population (W, small symbols). If local selection is weak, $s \lesssim 0.143$, runaway selection results in the fixation of the preferred trait T_1 in both populations. If selection is stronger, reinforcement of premating isolation occurs. (b) A similar transition can be observed if the transition probability, q, that determines the costs of mating preference is varied across $q \approx 0.925$. If costs are too high, $q \lesssim 0.654$, P_1 does not spread in either population. Parameters: (a) $q = 0.95$; (b) $s = 0.1$; both graphs: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.01$, $p_1 = 0.9$, $r = 0.5$, $t = 0.9$.

ment to occur, $s \gtrsim 0.143$. Then, the strength of reinforcement is largely independent of the strength of selection. If p_1 is below the threshold, a Fisherian runaway drives T_1 to fixation in both populations because $p_1 = 0.9 > s$ (compare the "rule of thumb" for the critical strength of mating preference, Eq. (5.32)).

A similar transition from runaway to reinforcement occurs if the transition probability crosses a threshold, $q \approx 0.925$ (see Fig. 5.9b). For small costs of mating preference (i.e., if q is above this threshold), runaway selection results in the fixation of T_1 in both populations, but below the threshold reinforcement occurs. A second threshold value exists, $q \approx 0.654$, below which P_1 can not spread in neither population. This effect was already observed in the mainland-island scenario that we analyzed in Section 5.3.1. To recapitulate, with increasing costs P_1 must generate stronger preference in order to spread in the uninfected population. Thus, with strength of mating preference being fixed, spread of P_1 is prevented if costs become too high.

Both critical influences of s and q on the transition between runaway and reinforcement can be understood from the fact that sexual and viability selection oppose each other in the infected population (whereas the two forces conjoin in the uninfected population). If sexual selection in favor of T_1 males outweighs viability selection in favor of T_2 males then spread of the trait T_1 is initiated, and a runaway spread of both T_1 and P_1 follows, to the effect of leveling nuclear divergence between the two populations. If, on the other hand, viability selection dominates sexual selection then spread of P_1 remains restricted to the uninfected population, and reinforcement yields premating isolation.

Increasing the selection coefficient directly strengthens viability selection while sexual selection remains unchanged, and thus the transition from a Fisherian runaway to reinforcement is easily understood. The key to understand why a similar shift occurs on decreasing the transition probability, is the fact that sexual selection is relaxed when costs become greater. To see this, consider a perfectly discriminating P_1 female, i.e., $p_1 = 1$. If transition into new mating rounds is certain, $q = 1$, then this female will eventually mate with a T_1 male even if it encounters nearly exclusively T_2 males. Thus, T_1 males perceive the full preference by P_1 females, and sexual selection is strong. If, however, it holds that $q < 1$ then T_1 males do not benefit from the full preference by P_1 females because some of them will not mate at all, as they reject unwanted T_2 males and are eventually excluded from the next mating round. The same reasoning applies to less discriminating females with $p_1 < 1$, but the effect is less pronounced.

Thresholds. In Fig. 5.10, the parameter plane spanned by the transition probability and the mating preference strength is depicted. Three major regions can be distinguished. (i) In the white region, P_1 can not spread in neither population. But in the gray shaded regions, P_1 does spread. (ii) In the region shaded in light gray, the spread is restricted to the uninfected population which results in premating isolation through reinforcement. (iii) In the dark shaded region, Fisherian runaway processes occur. In concordance with previous results, this region can be further subdivided: For weak mating preference, $p_1 \lesssim 0.17$, the P_1 allele becomes fixed in both populations; but stronger preference produces fixation of the sexually favored trait in both populations. In both cases, nuclear differences at the preference and

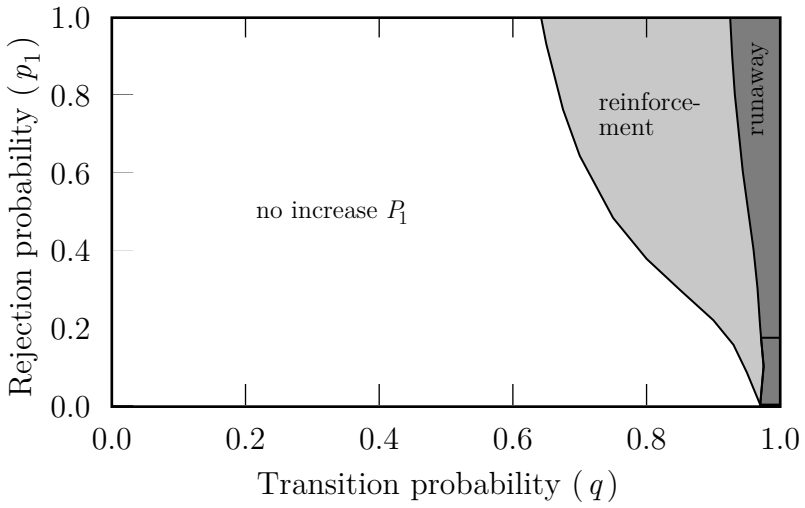


Figure 5.10: Thresholds of reinforcement. The plane spanned by the two parameters for the rejection probability of P_1 females, p_1, and for the transition probability, q, consists of three major regions. If q is too small, P_1 can not spread (white region). Provided however that q is large enough, the mating preference allele P_1 spreads (gray shaded regions) if the rejection probability is above a threshold. Reinforcement occurs in the region shaded in light gray, yielding between-population nuclear divergence at both the trait and the mating preference locus. Runaway processes occur in the dark gray region which is subdivided into a lower and an upper part. In the former, the runaway drives the preferred trait T_1 to fixation, whereas in the latter, P_1 becomes fixed in both populations. Parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.01$, $r = 0.5$, $s = 0.1$, $t = 0.9$.

trait loci are in effect erased. It is interesting to note that the transition between runaway processes and reinforcement is essentially determined by the transition probability and largely independent of the strength of mating preference.

The size of the white region in which no spread of mating preference occurs, can be attributed to the fact that small transition probabilities reflect very unfavorable ecological conditions for being a choosy female. After all, in our model one mating partner (the first one) is available for all females, so that trading this "secure" male for a low chance to find another mating partner must be considered indeed very costly. It should rather be stressed that, if mating preference is strong (for instance, if the rejection probability is larger than 80%) then P_1 already spreads even if the probability to encounter a new mating partner is no more than approximately 65% (also compare Fig. 5.9b).

To summarize, our models show that reinforcement occurs under a broad range of parameters and is favored by strong cytoplasmic incompatibility. It is critical that costs of mating preference neither be too small (lest a runaway process be triggered) nor too large (for spread of P_1 be possible at all).

5.3.4 Gene flow reduction

The fact that reinforcement of reproductive isolation takes place under a broad range of parameter values makes no assertion as to how this translates into reduction of gene flow. In this section, we will quantify the strengths of the different

barriers that act in reproductively isolating the populations. We use the approach of numerically calculating effective migration of neutral marker alleles as described in Section 4.2.5 of the previous chapter. According to Eq. (4.1), the gene flow factors calculated this way can be directly converted into strengths of the isolation barriers. Three isolation barriers can be distinguished in our scenario, building on cytoplasmic as well as nuclear differences between the populations: divergent viability selection (DS), cytoplasmic incompatibility (CI), and female mating preferences (MP) which we also refer to as sexual selection.

Nosil et al. (2005) have drawn attention to the fact that viability selection in ecologically divergent populations is an often underrated aspect in speciation processes. They conducted a re-analysis of existing studies on reproductive barriers and found that such selection against immigrants often constitutes a major component of reproductive isolation in the focal study systems. This effect also shows in our model due to the divergent selection pressures acting at the T locus. In order to restrict the impact of this barrier on isolation, we generally kept selection coefficients small. This allows for a better evaluation of our focal barriers, cytoplasmic incompatibility and sexual selection.

For a particular set of model parameter values ($f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.002$, $r = 0.5$, $p_1 = 1$, $q = 0.9$, $s = 0.1$, $t = 0.99$), Tab. 5.2 shows equilibrium frequencies of T_1, P_1, and *Wolbachia*, and the resulting gene flow factors (in percentages) for divergent selection only, for cytoplasmic incompatibility only, for these two barriers together, and for all three isolation barriers, i.e., after reinforcement has taken place. Three observations can be made: (i) Gene flow is reduced in a strong asymmetrical manner due to the infection polymorphism. (ii) For all practical purposes, divergent selection and cytoplasmic incompatibility act independently, i.e., synergistic effects are negligibly small. For instance, consider migration from the infected into the uninfected population, and assume complete independence of the two barriers. Multiplying the percentages of gene flow that passes through the DS and CI barriers alone yields $83.9\% \cdot 10.4\% = 8.73\%$ gene flow. Thus, by comparing this number to

Reproductive isolation barrier	Equilibrium frequencies						Gene flow factor (%)	
	T_1		W		P_1			
	1	2	1	2	1	2	$1 \leftarrow 2$	$1 \rightarrow 2$
none	–	–	–	–	–	–	100.0	100.0
DS only	0.98	0.02	–	–	–	–	83.9	83.9
CI only	–	–	0.02	0.998	–	–	10.4	54.3
DS + CI	0.997	0.011	0.016	0.999	–	–	8.6	45.2
DS + CI + MP	0.999	0.008	0.016	0.999	0.948	0.002	3.2	30.7

Table 5.2: Gene flow factors of reproductive isolation barriers. Numerically calculated equilibrium frequencies of alleles T_1 and P_1 and of *Wolbachia (W)* in populations 1 and 2. Without isolation barrier, the gene flow factor (calculated as described in Section 4.2.5) is 100% for both directions of migration. Divergent selection (DS) at the trait locus and cytoplasmic incompatibility (CI) can function as isolation barriers on their own. If combined, only small synergistic effects are observed, they largely act independently. The spread of the mating preference (MP) allele P_1 further reinforces repoductive isolation. Parameters: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.002$, $r = 0.5$, $p_1 = 1$, $q = 0.9$, $s = 0.1$, $t = 0.99$.

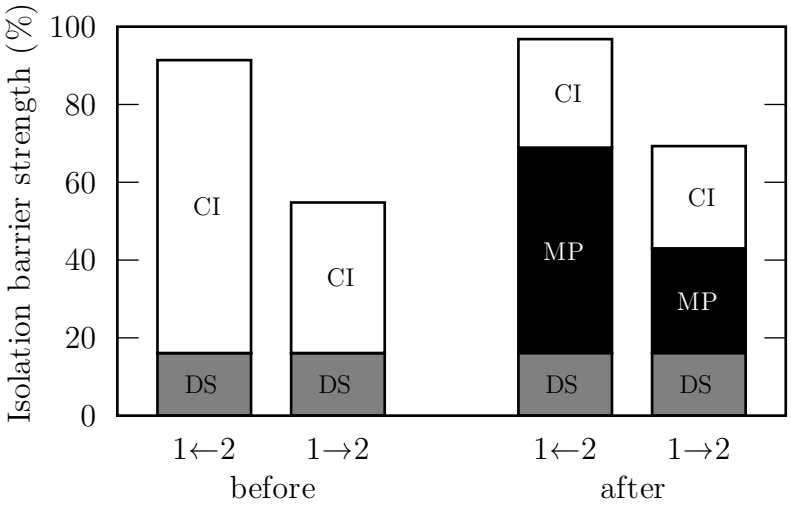


Figure 5.11: Relative strength of isolation barriers before and after reinforcement. In the life cycle of the modeled organisms, the sequential order in which reproductive isolation barriers take effect is: divergent selection (DS), mating preference (MP), and cytoplasmic incompatibility (CI). Evolution of the MP barrier reinforces isolation in the direction from the infected into the uninfected population ($1 \leftarrow 2$) stronger than in the opposite direction ($1 \rightarrow 2$). The relative contribution of the CI barrier diminishes because it only acts on the gene flow that passes the MP barrier. Absolute strength of isolation is given by the total height of a bar. Parameters in the underlying simulations: $f = 0.1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.002$, $p_1 = 1$, $r = 0.5$, $q = 0.9$, $s = 0.1$, $t = 0.99$.

the numerically calculated 8.6% listed in Tab. 5.2, it can be concluded that synergistic effects only account for a further reduction by 0.13%. (iii) Reinforcement due to the spread of the P_1 allele in the uninfected population increases isolation in both directions of migration, albeit in an asymmetrical fashion (more details below).

Based on the sequential order in which the isolation barriers act on gene flow, the relative contributions of the three barriers can be computed (relative to the total isolation in a given scenario and direction of migration). Fig. 5.11 shows a bar chart of the relative strengths of the isolation barriers before and after reinforcement, converted from the gene flow factors of Tab. 5.2. Absolute strength of isolation in each direction is given by the total height of a bar. The strength (absolute as well as relative) of the barrier due to divergent trait selection is not affected by reinforcement because in our model, viability selection occurs after migration but before mating which includes sexual selection (see Section 5.2). For a selection coefficient of $s = 0.1$ and migration occurring at a rate $m = 0.002$, the relative strength of the DS barrier is 16.1% in both directions of migration. Before reinforcement, the relative contributions of the CI barrier to gene flow reduction are 75.3% (for migration from the infected into the uninfected population) and 38.7% (for the opposite directions). After reinforcement, however, CI barrier's relative strength has decreased to 27.9% and 26.3%, respectively. Mating preferences contribute relative isolation strengths of 52.8% and 26.9%, respectively. Thus, reinforcement is stronger (roughly twice as strong) for the direction that is already impaired more due to the asymmetrical nature of unidirectional CI.

Gene flow reduction is conferred by different sexes of migrants in the two direc-

tions. In the uninfected population, gene flow through *male* immigrants is impeded because these are less attractive to resident females whereas female immigrants are unaffected by sexual selection. In a reverse manner, in the infected population gene flow through *female* immigrants is reduced because these keep rejecting local males due to their sexual preferences whereas male immigrants are not subject to sexual selection.

From the previous Chapter 4, we know that both the transmission rate of *Wolbachia* and the level of CI are of great importance for the reduction in gene flow. The closer these two parameters are to one, the more does the gene flow from the infected to the uninfected population tend against zero (compare Eq. (4.17)). Importantly, reinforcement still occurs if gene flow is already completely extinguished. This is because it is even the more beneficial for an uninfected female to avoid matings if they do not yield viable offspring at all.

Gene flow from the uninfected into the infected population, on the other hand, can at best be reduced to 50% due to CI (see Eq. (4.10) in Section 4.2.1). However, evolution of female mating preferences in the uninfected population can substantially reinforce isolation in this direction of migration as we have shown above.

In the scenario from Section 5.3.1, migration only occurs from an infected mainland to an uninfected island. Under these circumstances, the following gene flow factors for the otherwise same set of parameters as in Tab. 5.2 can be calculated through numerical simulations. Viability selection alone results in gene flow reduction to 83.6%, CI alone is responsible for a reduction to 10.5%, the two combined result in a gene flow factor of 8.6%, and after reinforcement has driven P_1 to an equilibrium frequency of 0.949 in the uninfected population, total gene flow is down to 3.2%. Thus, nearly the same reductions in gene flow are attained as in the scenario with two-way migration.

It is important to acknowledge that gene flow will never be completely shut down in these reinforcement scenarios even under most favorable conditions because migrating females with the wildtype preference allele P_0 will always ensure some degree of genetic exchange. Nevertheless, we conclude that reinforcement of premating isolation due to a postzygotic CI barrier does substantially reduce gene flow between populations. The effect is asymmetric in that gene flow is reduced stronger for migration from infected into uninfected populations than in the other direction. The relative importance of the CI barrier is diminished after establishment of sexual selection.

5.4 Discussion

We have analyzed in this chapter whether a stable coexistence of parapatric uninfected and infected host populations selects for divergence at a locus for female mating preference such that the postzygotic CI barrier reinforces premating isolation. Our main findings are: (i) Premating isolation readily evolves between an uninfected island population and an infected mainland due to the spread of a female mating preference mutant on the island. (ii) In two-way migration scenarios, a stable infection polymorphism can cause runaway selection (Fisher, 1930) if costs of female mating preference are small. Such Fisherian runaway processes lead to either the fixation of female mating preference or of the preferred trait in both

populations. In both cases, premating isolation is not established. (iii) If costs of mating preference are large enough (but not too large), unidirectional CI can reinforce the evolution of premating isolation even if migration occurs in both directions. (iv) Reinforcement results in decreased gene flow in both directions of migration but predominantly affects migration from infected to uninfected populations. This leads to a diminished relative contribution of the CI barrier to reproductive isolation between populations. (v) In the scenarios analyzed here, spread of female mating preference is a slow process.

The runaway processes observed in our model are equivalent to what Telschow et al. (2005a) reported in a theoretical study that compared nuclear incompatibilities with *bidirectional* CI in terms of reinforcement. Telschow and colleagues cautioned that because runaway sexual selection can destroy male trait polymorphism, selection for mating preference does not necessarily reinforce premating isolation. Therefore, in order to detect reinforcement it does not suffice to analyze invasion dynamics of mutants at the mating preference locus (e.g., Servedio, 2000). Reinforcement can only be claimed if divergence is maintained until equilibrium is reached. Our simulations confirm these conclusions. It should be noted, however, that one possible way out of this problem is that female mating preferences are assortative rather than for fixed male traits (e.g., Proulx and Servedio, 2009). With assortative mating, a mutant allele could spread in both populations and in each population become associated with the respective locally adaptive trait. This would result in premating isolation (but without divergence at the preference locus).

Leaving the CI polymorphism aside for a moment, we can ask: Why does divergent trait selection not suffice for female mating preference to spread in a two-population system? The key point in understanding this feature is that producing fitter and more attractive sons is only beneficial for choosy females as long as the trait itself spreads. In mathematical models, Lande (1981) and Kirkpatrick (1982) derived the existence of a neutrally stable line of equilibria with respect to trait and preference allele frequencies. They showed that female mating preference evolves as a correlated response to the spread of the sexually favored trait. In turn, this means that if the trait is already close to fixation, as is the case in our models with local selection at the trait locus, then mate discrimination will generally not spread.

However, Bulmer (1989) demonstrated that the above models are structurally unstable if mutation is taken into account. More importantly, direct selection at the mating preference locus also changes the system's behavior. The dominant force in determining the direction and extent of sexual selection is likely to be selection acting directly on the female preference. In our model, migration and the direct benefit that discriminating females gain from CI avoidance, have similar effects. In general, it is the female preference (and not the male trait) that will find its optimal level under forces of mutation, direct selection, and genetic drift, dragging the male trait with it along the line of equilibria.

There are several reasons why female mate discrimination could be associated with costs. In general, any cause will do that reduces the probability of a female —after having rejected a possible mating partner— to survive long enough to find another, potentially better suited partner. This implies that the ecology of the organisms under study, and especially the mating system (Choe and Crespi, 1997) largely determines the costs of female mating preferences. For instance, if mating

usually takes place in lek mating arenas, many males will be readily available to females but aggressive male mating behavior might still induce costs. On the other hand, if mating does not occur at male aggregation sites, increased rates of predation will likely pose a major risk when rejecting a male. In fact, it seems difficult to conceive of an ecology that —in one way or another— would *not* make female mating preferences costly. The amount of costs, however, might vary considerably between ecologies and mating systems.

Throughout this chapter, we have been characterizing the spread of the mating preference allele P_1 and the preferred trait T_1 in the infected host population as a Fisherian runaway process because part of the driving force behind this spread is the combined advantage that sexual selection confers on T_1 males and the advantage to produce more attractive sons for P_1 females. However, we do not observe the "ever-increasing speed" that characterizes the processes imagined by Fisher (1930, see also Section 2.2.2). Quite the contrary, the dynamics described in this chapter are often remarkably slow. There are two major reasons for this lack of speed: First, the runaway must proceed against viability selection which can be considered the "counterselection" that (Fisher, 1930) also predicted to eventually stop runaway processes. In our scenarios, however, this counterselection does not abruptly kick in at the end of the runaway process but rather perpetually slows it down. Second, sexual selection also builds up due to the spread of female preference in the uninfected population and continuous immigration of these choosy females into the infected population. The spread within the uninfected population is a slow process by itself, and thus also immigration only slowly increases.

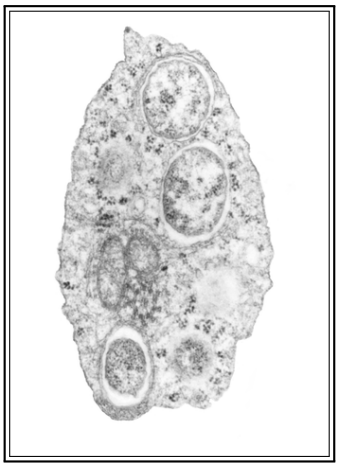
Many traits that influence sexual selection are complex behaviors. In *Drosophila*, for instance, male courtship songs are crucial (e.g., Ritchie et al., 1999), but at the same time cuticular hydrocarbons play an important role by functioning as contact pheromones (Gleason et al., 2005). This can strongly influence mate recognition because the cuticular hydrocarbon composition varies both between and within *Drosophila* species (Jallon and David, 1987, e.g.,). Thus, the abstract model of sexual selection that we adopted in this chapter, i.e., a single male trait that is used as a mating cue by females, can be considered to reflect one aspect of a more complex mating system, and therefore is applicable to a broad range of systems.

Our model scenario involves a full allopatric phase during which divergence at the trait locus evolves. The mutant at the preference locus is introduced after secondary contact is established. An alternative model setting would allow a certain amount of migration from the start. It is important to remark that the full allopatry at the beginning is not a necessary condition for reproductive isolation to evolve. A modification of our model with migration from the start does not change the results qualitatively as long as migration is below the critical migration rate.

In this study, we analyzed the question whether a single *Wolbachia* strain causing unidirectional CI can select for local adaptation and premating isolation. Previously, it was shown that two *Wolbachia* strains causing bidirectional CI promote local adaptation and select for premating isolation under a broad variety of conditions (Telschow et al., 2005a). This is mainly because bidirectional CI can persist up to high critical migration rates of over 15% per generation (Telschow et al., 2005b), and higher migration rates produce stronger selection pressures to avoid incompatible matings. As shown in this thesis, critical migration rates for unidirec-

tional CI are much lower, resulting in comparatively weak selection for premating isolation. These results suggest that once bidirectional CI is established in a system, its impact on host speciation is much stronger than that of unidirectional CI. Nevertheless, the full significance of unidirectional CI for host speciation might be larger than that of bidirectional CI. First of all, unidirectional CI is much more common in nature because it requires that only one population is infected whereas bidirectional CI requires acquisition of two incompatible strains of *Wolbachia*. Furthermore, in case of bidirectional CI, one *Wolbachia* strain is inevitably eliminated if postzygotic isolation is destroyed. Postzygotic isolation can be reestablished only if the system becomes infected again with a second *Wolbachia* strain, a rather unlikely event. In the unidirectional CI scenario, however, reestablishment of postzygotic isolation involves only the loss of the infection in an island population, an event which is much more likely to reoccur. In fact, as discussed already in the previous Chapter 4, recurrent establishment of peripheral isolate populations from a central source population could produce the circumstances for repeated opportunities for local adaptation and speciation. Our results suggest that when these conditions occur, those peripheral populations that have lost their infections are more likely to maintain locally adapted alleles and to evolve reinforcement of mate discrimination.

In summary, our findings support the view that *Wolbachia*-induced unidirectional CI could be a factor in host divergence and speciation, but only under certain conditions. If the infection polymorphism is stable, CI acts as a postzygotic isolation mechanism by reducing gene flow between populations, selects for local adaptation, and reinforces premating isolation. Gene flow is reduced asymmetrically by CI, and reinforcement strengthens this pattern. The full impact of unidirectional CI on host speciation in nature will likely involve more complex population structures than modeled here (see also Chapter 6). The results on the critical migration rates suggest, however, that *Wolbachia* infections might occur in mosaic patterns with infected and uninfected patches close by. We have shown here that under these circumstances unidirectional CI selects for female mating preferences in uninfected patches facing migration from infected patches.



Chapter 6

Mycophagous *Drosophila* in North America: applying the model

Two sibling species of mycophagous *Drosophila* show overlapping geographic ranges in North America and also occur in microsympatry. Mating experiments have revealed patterns of reproductive character displacement consistent with the expected outcome of reinforcement processes driven by cytoplasmic incompatibility.

In this chapter, we adjust our model to accommodate the empirical setting of this real-world speciation process. We present results from numerical simulations indicating that *Wolbachia*-induced cytoplasmic incompatibility can select for the observed patterns of sexual selection and thus reinforce premating isolation but only under stringent conditions. Female mating preference must incur costs, and trait and preference loci must be physically linked as to reduce recombination. Based on our findings, we suggest that some characteristic of the system that stabilizes the *Wolbachia* infection polymorphism might have been overlooked so far.

> The foundations of population genetics were laid chiefly by mathematical deduction from basic premises contained in the works of Mendel and Morgan and their followers. Haldane, Wright, and Fisher are the pioneers of population genetics whose main research equipment was paper and ink rather than microscopes, experimental fields, *Drosophila* bottles, or mouse cages. Theirs is theoretical biology at its best, and it has provided a guiding light for rigorous quantitative experimentation and observation.
>
> Dobzhansky (1955, pp. 13–14)

6.1 Introduction

In evolutionary biology, just like in many other areas of scientific research, computer simulation models have long become an essential supplement to analytical models. However, as a refinement of intuition it is essential that such theory be constantly checked against empirical data to remain anchored to reality. In this chapter, we aim to apply the models developed in the previous chapters to a real-world speciation process, the case of two mycophagous *Drosophila* species in North America, *D. subquinaria* and *D. recens* (Shoemaker et al., 1999; Jaenike et al., 2006; Jaenike, 2007).

D. subquinaria and *D. recens* are closely related sibling species of the quinaria species group (compare the phylogeny in Fig. 6.1). A screening for *Wolbachia* in the quinaria group (Werren and Jaenike, 1995) revealed that two species were infected: *D. recens* and *D. orientacea*. Although *D. subquinaria* was not included in this study, Shoemaker et al. (1999) determined that this species does not harbor a *Wolbachia* infection.

Most of the quinaria subspecies are mycophagous, and many of the sibling species tested are not only sympatric in the wild but are also found on the same individual mushrooms suggesting ample ecological opportunity for hybridization and horizontal transfer. This applies to both *D. subquinaria* and *D. recens*. Mushrooms serve as larval food resources, and adults of these species feed, court, mate, and oviposit on mushrooms (Lacy, 1984; Jaenike et al., 2006). The entire life cycle takes about fourteen days under normal conditions (Lacy, 1984).

By rearing larvae of *D. recens* on tetracycline-treated mushrooms, Werren and Jaenike (1995) obtained an uninfected strain of *D. recens* and demonstrated that *Wolbachia* induces CI in this host. The CI phenotype was corroborated by Shoemaker et al. (1999) as it was evident in crosses between *D. subquinaria* females and *D. recens* males. An additional form of reproductive isolation between *D. subquinaria* and *D. recens* is present: In concordance with Haldane's rule, female hybrids are fully viable and fertile whereas male hybrids are viable but completely sterile.

In the remainder of this introductory section, we will focus on the study by Jaenike et al. (2006), and relate the main findings and conclusions. Fig. 6.2 is adapted from this study and shows that the geographical ranges of *D. subquinaria* (horizontal hatching) and *D. recens* (vertical hatching) overlap in central Canada.

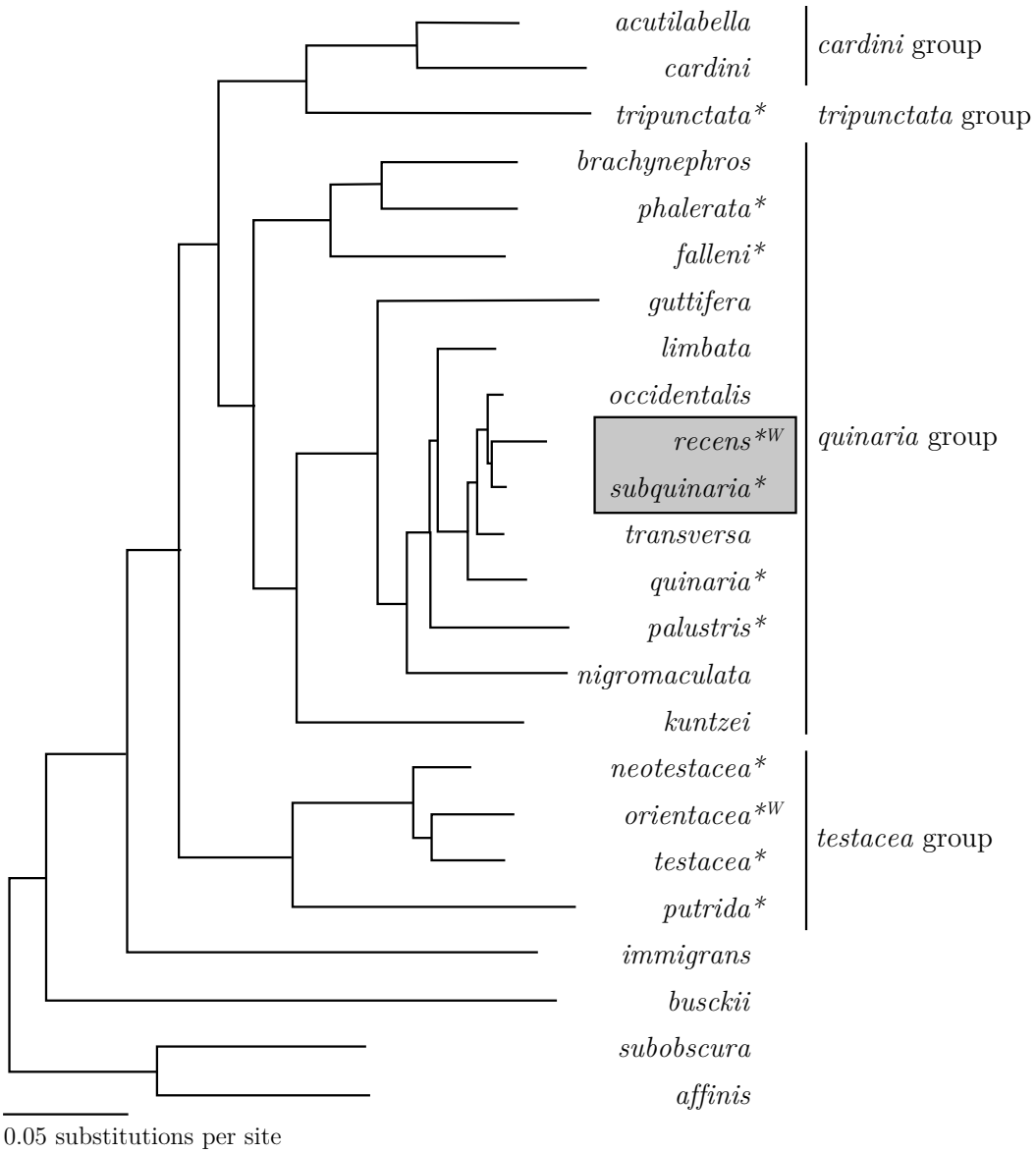


Figure 6.1: Phylogeny of the *Drosophila quinaria* species complex. Maximum likelihood tree using sequences of mitochondrial cytochrome oxidases I, II and III (adopted from Perlman and Jaenike, 2003; Perlman et al., 2003). The gray shaded box highlights the two sibling species that are the subject of study in this chapter. Asterisks indicate species screened for *Wolbachia* infections (Werren and Jaenike, 1995; Shoemaker et al., 1999), species that were found to be infected are marked by a superscript *W*.

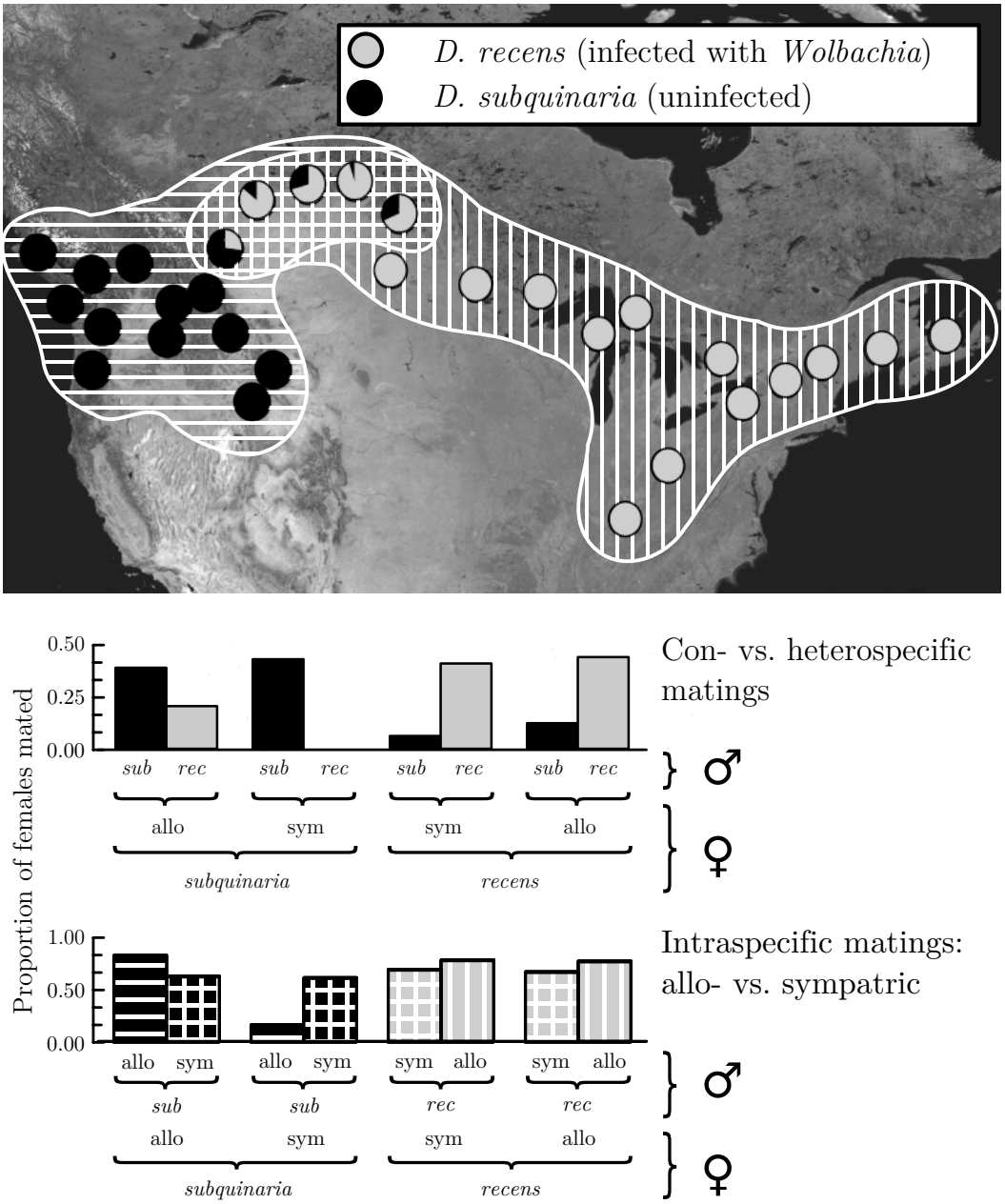


Figure 6.2: Asymmetric reinforcement in North American *Drosophila*. The geographic ranges of uninfected *D. subquinaria* and *Wolbachia*-infected *D. recens* overlap in central Canada. Mating experiments revealed asymmetric reproductive character displacement: *D. subquinaria* females sympatric with *D. recens* showed strong between-species and within-species mating preferences. For more details, see main text. The map with the collection sites is taken from Jaenike et al. (2006). We recompiled the bar charts from data from the same study. Abbreviations: sym–sympatric, allo–allopatric, *sub*–*D. subquinaria*, *rec*–*D. recens*.

Circles represent collection sites of *Drosophila* flies with black denoting proportions of *D. subquinaria* and gray denoting proportions of *D. recens*. Females of both species show mate discrimination against heterospecific males, but within the zone

of sympatry which is at least 1,200 km wide, *D. subquinaria* females exhibit stronger mating preference than females of *D. recens*. Sympatric *D. subquinaria* females not only discriminate against *D. recens* males but also against conspecific *D. subquinaria* males from allopatric populations. This pattern of greater prezygotic isolation in sympatry relative to allopatry is one of the main "signatures" of reinforcement and has been termed reproductive character displacement (Servedio and Noor, 2003).

The bar charts[1] in Fig. 6.2 demonstrate the reproductive character displacement. Bars give proportions of mated females in the two types of mating experiments that Jaenike et al. (2006) conducted. In "mass population" trials, matings within populations of 200 females and 100 males were recorded, whereas in "no-choice" trials, individual females were presented with single males. The upper bar chart in Fig. 6.2 is compiled from mass population experiments, whereas the lower chart is based on no-choice trials (the differing y-axis scales of the two charts derive from this difference; for more details on the trials, see the *Materials and Methods* section in Jaenike et al., 2006). Bars are ordered in groups of two, based on geographic origin and species of the female mating partner, i.e., from left to right: allopatric *D. subquinaria*, sympatric *D. subquinaria*, sympatric *D. recens*, and allopatric *D. recens*. The color of a bar corresponds to the species and the hatching to the geographic origin of the mated male.

Patterns of variation at the mitochondrial *Cytochrome Oxidase I (COI)* locus reveal little genetic differentiation between sympatric and many allopatric populations within each species. This suggests that reproductive character displacement has evolved in the face of considerable gene flow and is likely the result of selection rather than genetic isolation among populations. The collected samples included three (uninfected) *D. subquinaria* individuals that had a *D. recens* mtDNA haplotype indicating that hybridization has occurred in the past. By contrast, no *D. recens* individuals with mtDNA from *D. subquinaria* were found. This is expected because uninfected *D. subquinaria* females suffer from CI when immigrating into *D. recens* populations so that introgression of *D. subquinaria* mtDNA haplotypes is strongly impeded in this direction of migration. Jaenike et al. (2006) concluded that the pattern of female mating preferences is in accordance with models of reinforcement due to the effects of *Wolbachia*-induced CI upon hybridization.

In the present study, we inspect whether this conclusion is supported by our models of speciation via reinforcement developed in the previous chapters.

6.2 Model

We rely on the basic model developed in the previous chapter, and we use the same mathematical formalization (compare Section 5.2). To account for the focal study system of North American *D. subquinaria* and *D. recens*, we mainly add two features: We include a simple genetic model for the hybrid male sterility (HMS) found in matings between the two sibling species, and we map the more complex population structure (compare Fig. 6.2) to a four-population stepping stone model, with each species consisting of two populations. Here, we verbally describe the general model which we formalize mathematically in Appendix C. Note that in

[1] Charts are recompiled from data from Jaenike et al. (2006).

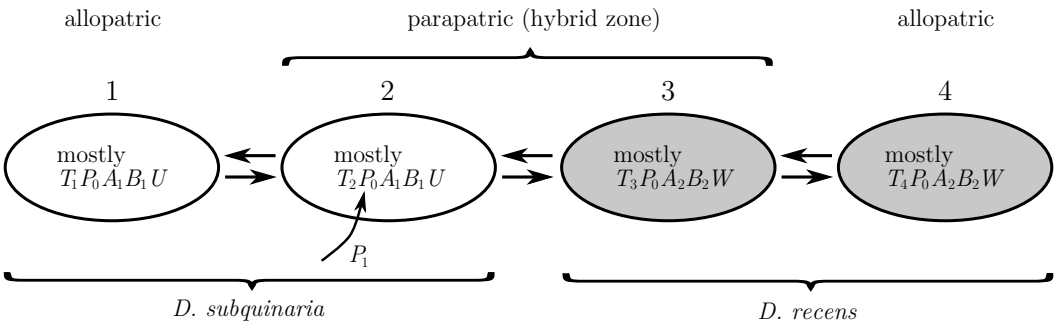


Figure 6.3: Stepping stone population structure and general model scenario. The geographic ranges of uninfected (U) *D. subquinaria* and *Wolbachia*-infected (W, populations shaded in gray) *D. recens* are mapped to four populations (numbered from left to right) in a stepping stone geometry with neighboring populations connected by two-way migration. The outermost populations are called allopatric, the two inner populations comprise the parapatric hybrid zone. Local viability selection favors a different trait, T_i, in each population i. At two background loci, A and B, different alleles have become fixed in the two leftmost populations (A_1B_1) than in the two rightmost populations (A_2B_2). Hybrid male sterility is caused by epistatic interactions between these loci. Wildtype females with the allele P_0 at a locus for female mating preference mate randomly. A mutant allele, P_1, which expresses a preference for the trait T_2, is introduced in the parapatric *D. subquinaria* population.

the results section, we will not always use the full model but also analyze certain features of the study system in "subsets" of the model.

First, we briefly recapitulate the basic features of the model developed in the last chapter. Host organisms are haploid but reproduce sexually in a short diploid phase. Generations are discrete and non-overlapping, and the organisms' life cycle consists of four stages: migration, viability selection, sexual selection, and reproduction. Populations are considered to be infinitely large and perfectly mixing. Gene exchange between populations occurs through migration. Organisms are fully characterized by their nucleocytotype.

The nuclear genotype in this chapter's model is given by the allele composition at four loci. The first two loci are the same as in the previous chapter: a locus for a phenotypic trait, T, and a locus for female mating preference, P. The other two comprise background loci, A and B, that are involved in hybrid dysfunctions (more detail below). The cytotype is equivalent to an organism's infection status (either infected with *Wolbachia* or uninfected). *Wolbachia* dynamics are determined by the fecundity costs that infected females incur, by the rate of maternal transmission, and by the level of cytoplasmic incompatibility.

The real population structure of the two *Drosophila* species under study is not known, but the collection sites relate a complex picture (compare Fig. 6.2). It would be hopeless to expect any general conclusions if one was to model this structure in its full complexity. Instead, we simply allow for within species reproductive character displacement by extending our model to four populations that are arranged in a stepping-stone geometry (see Fig. 6.3). All neighboring populations are linked by bidirectional migration; the two leftmost populations comprise the geographic range of uninfected *D. subquinaria*, and the two rightmost populations the range

of *Wolbachia*-infected *D. recens*. We will refer to the two inner populations as "parapatric" and the two outer populations as "allopatric", reflecting the degree of contact with respect to the other species. The two inner populations then can be considered the hybrid or contact zone. We want to remark that the use of the term "species" in our model is merely a matter of conveniently referring to sets of populations (the two leftmost or the two rightmost, respectively).

In each of the four populations, a different allele at the trait locus, T, is favored by local viability selection. Numbering the populations from left to right, T_1 is adaptive in population 1, T_2 in population 2, T_3 in population 3, and T_4 in population 4.

Furthermore, we assume that the two species have diverged in allopatry (i.e., when the innermost migration links between populations 2 and 3 were interrupted, presumably due to one of the glaciation periods in the late Pleistocene) at the two background loci, A and B. From an ancestral set of alleles, A_0B_0, a new set A_1B_1 has spread through the two leftmost populations, whereas in the two rightmost populations, A_2B_2 has become fixed. This is similar to the modeling of nuclear Dobzhansky-Muller incompatibilities (Dobzhansky, 1937; Muller, 1942) in several theoretical reinforcement studies (e.g., Kirkpatrick and Servedio, 1999; Servedio, 2000; Telschow et al., 2002a). However, we assume that a hybrid genetic background, i.e., the two allele combinations A_1B_2 and A_2B_1, does not yield viability defects but instead results in complete sterility of males.

Our model of hybrid male sterility is consistent with the faster-male theory (Wu and Davis, 1993) which is based on the assumption that incompatibilities that afflict heterogametic hybrids are more common than those afflicting homogametic hybrids (for a review, see Coyne and Orr, 2004). Such difference may arise if (i) spermatogenesis is an inherently sensitive process that is easily perturbed in hybrids or if (ii) sexual selection causes faster evolution of male- than female-expressed genes. While the faster-male theory cannot explain Haldane's rule in general as it aims at sterility only and —more importantly— is only applicable to male heterogametic species (Coyne and Orr, 2004), there is now good evidence that it underlies HMS in *Drosophila* (e.g., True et al., 1996; Tao et al., 2001; Ranz et al., 2003; Zhang et al., 2004). In fact, male genitalia are about the only diagnostic differences between *D. subquinaria* and *D. recens* males (Jaenike et al., 2006). Note that for simplicity, we assume purely autosomal incompatibilities that only take effect in males.

At the locus for female mating preference, there are two possible alleles, the wildtype, P_0, and a mutant, P_1. Females carrying P_0 do not show mating preference and accept all males indiscriminately. P_1 females, however, reject with a certain probability males not showing the trait T_2 that is adaptive in population 2. Female mating preference may be costly, e.g., due to energetic costs of sampling or refusing to mate and lost breeding opportunities caused by rejecting potential mates.

In our model scenario, we assume for each population a fixed starting nucleocytotype as depicted in Fig. 6.3: In each population, the locally adaptive trait is fixed. All host organisms in all populations have the wildtype mating preference allele at the beginning. In the two leftmost populations (comprising the species range of *D. subquinaria*), background alleles A_1B_1 are fixed, and all organisms are uninfected. In the two rightmost populations (comprising the range of *D. recens*), all organisms have alleles A_2B_2 and harbor *Wolbachia*. Following the onset of migration and the establishment of an equilibrium state, the mutant allele P_1 is introduced with low

frequency in population 2.

A precise mathematical formalization of the verbal model description is given in Appendix C. Table 6.1 presents an overview of the parameters and gene loci. The model is by far too complex to be treated analytically. We have therefore reverted to numerical simulations. To narrow down the parameter space, we derive parameter estimates from the scientific literature on *D. recens* and *D. subquinaria* in the following section.

The main objective of our analysis is to corroborate or invalidate the hypothesis that the infection polymorphism can reinforce premating isolation and produce the pattern of reproductive character displacement evident from the mating experiments by Jaenike et al. (2006).

6.2.1 Parameter estimates

From the focal study by Jaenike et al. (2006), we can infer values for fecundity reduction, CI level, and transmission rate. The *Wolbachia* infection of *D. recens* does not seem to affect female fecundity, i.e., $f = 0$. The level of cytoplasmic incom-

Symbol	Value	Description
f	0	Female fecundity costs of a *Wolbachia* infection
h	1	Level of hybrid male sterility
l_{CI}	0.9	Level of cytoplasmic incompatibility; CI level
$m_{i,j}$	0.005	Rate of migration from population j to population i
p_1	$0 \leq r \leq 1$	Rejection probability of mating preference allele P_1
q	$\lesssim 1$	Transition probability between mating rounds
r	$0 \leq r \leq 0.5$	Recombination rate between nuclear loci T and P
s	0.1	Coefficient of divergent viability selection at the T locus
t	0.87	Transmission rate of *Wolbachia*
T_1, T_2, T_3, T_4		Alleles at nuclear trait locus T
P_0, P_1		Alleles at nuclear locus P for female mating preference
A_1, A_2		Alleles at nuclear background locus A
B_1, B_2		Alleles at nuclear background locus B
U, W		Cytotypes, or infection statuses; U–uninfected, W–infected with *Wolbachia*

Table 6.1: Glossary of model notation. Overview of the parameters and symbols used in the model of this chapter. Parameter values or ranges are also given. The background loci A and B are involved in hybrid male sterility. For more details, see main text but also compare Tab. 5.1.

patibility has been measured to be $l_{\mathrm{CI}} \approx 0.9$. Although Jaenike et al. (2006) found perfect transmission of *Wolbachia* in their laboratory studies, this might not hold true in the field (Turelli and Hoffmann, 1995). According to our modeling of a single panmictic host population (see Section 3.2.1), the measured infection frequency of 98% (Jaenike et al., 2006) can be obtained by assuming imperfect *Wolbachia* transmission with $t \approx 0.87$ (cp. Eq. (3.33b)). Furthermore, we can calculate the critical migration rate for the spread of *Wolbachia* based on these parameter estimates. For symmetrical migration, Eq. (3.40) yields $m_{\mathrm{crit}} \approx \min\{0.006, 0.148\}$. As we restrict our analysis to symmetric migration between all "stepping stones", we will therefore only use migration rates, $m < 0.006$.

Hybrid males from interspecific crosses invariably lack motile sperm (Shoemaker et al., 1999) and are —in concordance with Haldane's rule— fully sterile so that we can infer $h = 1$. This hybrid male sterility and the absence of hybrid breakdown (Shoemaker et al., 1999) indicate an early stage in speciation. Based on mtDNA nucleotide diversity, Shoemaker et al. (1999) estimated that the two sibling species diverged approximately 0.6 million years ago. Within *D. recens* mtDNA haplotype diversity suggests that *Wolbachia* may have swept this species about 50,000 years ago. Secondary contact has presumably taken place 10,000 to 12,000 years ago with the retreat of the Wisconsin glacier. Mushroom-feeding quinaria flies pass the winter as adults in reproductive diapause (Jaenike, 1992), but they breed continuously from May through October (Lacy, 1983). As stated above, the normal generation time of *D. recens* and *D. subquinaria* is fourteen days (Lacy, 1984). Thus, conservatively assuming an average generation time of three weeks, and assuming 26 weeks of possible reproduction per year, the roughly 10,000 years since secondary contact can be equated to about 90,000 generations. We will consider this the maximal time frame for the presumable spread of female mating preferences and establishment of reproductive character displacement. If our model yields reinforcement but dynamics are significantly slower than this time span then we will have to conclude that CI-driven reinforcement is unlikely the cause for the observed pattern of mating preferences.

Sympatric *D. subquinaria* females fully reject *D. recens* males. Therefore, the rejection probability of the mutant allele at the locus for female mating preference, P_1, is one: $p_1 = 1$. However, we will also consider different rejection probabilities to investigate whether a stepwise increase in preference strength, starting with low strength, is evolutionarily feasible.

We conclude our parameter estimates with the statement that the two mycophagous species of *D. recens* and *D. subquinaria* often aggregate on the same mushrooms. Therefore, it seems most realistic to assume small costs of mate rejections as other mating partners will likely be available so that the transition probability, q, between mating rounds might not deviate much from one.

This leaves two parameters of our model without estimate, the coefficient of selection acting on the trait locus, s, and the rate of recombination between trait and preference loci, r. In our analysis, we will mostly stick to weak selection, $s = 0.1$, as we have done in the previous chapter, and we will allow the recombination rate to take values between zero and one half, $0 \leq r \leq 0.5$.

6.3 Results

The results section consists of three parts. First, we present an infection stability analysis (compare Chapter 3) that takes into account the sterility of hybrid males. Then, we present equilibrium states following introduction of the preference mutant, P_1, that are representative for the possible outcomes of the system dynamics. In this part, we exclude the two background loci in order to focus on CI as a driving force for the evolution of female mating preferences. Finally, we demonstrate that in the full model, reinforcement can occur under rather restrictive parameter conditions, and we show how gene flow is affected in this case.

6.3.1 Infection polymorphism stability

In short, the stability of a *Wolbachia* infection polymorphism is practically unaffected by the presence of HMS. In Fig. 6.4, critical migration rates for the two-population model with symmetric bidirectional migration are plotted as a function of the CI level. In Chapter 3, we have used analytical formulas for mainland-island scenarios to estimate critical migration rates for the case of two-way migration. In this chapter, fecundity costs are always assumed to be absent, $f = 0$, so that Eq. (3.40) applies. For $t = 0.87$, the numerically calculated critical migration rate is plotted as a dashed line in Fig. 6.4. With HMS, the critical migration rate is very slightly increased (solid line). For all practical purposes, synergistic effects between cytoplasmic incompatibility and hybrid male sterility do not occur.

It has been speculated that nuclear incompatibilities could significantly stabilize infection patterns (e.g., Telschow et al., 2007). In fact, it has been shown in

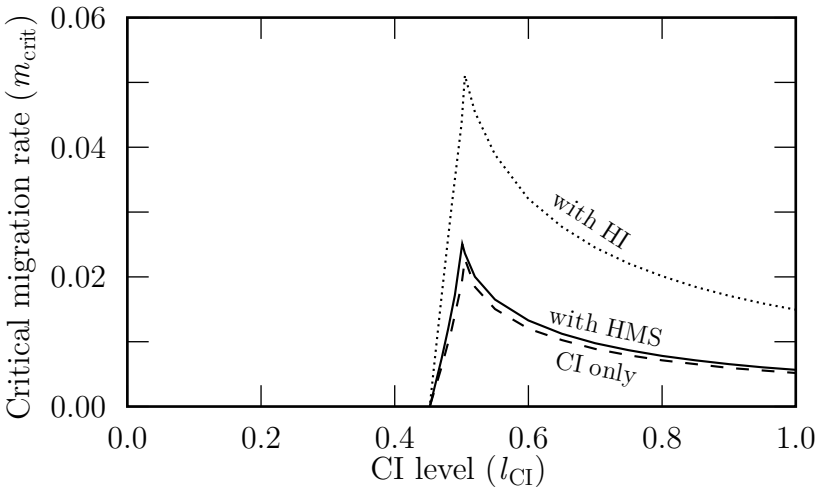


Figure 6.4: Effects of nuclear incompatibilities on infection stability. The stability of the *Wolbachia* infection polymorphism, as measured by the critical migration rate (see Chapter 3), is only very slightly increased by the presence of hybrid male sterility (HMS, solid line) if compared to the model with cytoplasmic incompatibility only (CI, dashed line; compare Eq. (3.40) and Fig. 3.10a). By contrast, complete hybrid inviability (HI, dotted line) would significantly boost the stability. All curves have been derived from numerical simulations. Parameters: $f = 0$, $h = 1$ (both HMS and HI), $t = 0.87$.

a diploid model that the spread of *Wolbachia* onto an uninfected island (compare Section 3.2.3) can be significantly impeded by dominant lethal nuclear incompatibilities (Hilgenböcker, 2009). However, if nuclear incompatibilities obey Haldane's rule then infection polymorphisms are only stabilized in species with heterogametic females, i.e., the *Wolbachia*-transmitting sex must be affected by the incompatibilities (Hilgenböcker, 2009). Our haploid model is consistent with these findings. Replacing the male sterility effect of recombinant A_1B_2 and A_2B_1 with complete inviability in our model equations is simple (see Appendix C). The dotted curve in Fig. 6.4 depicts the increase of the critical migration rate if hybrids are completely inviable. However, the absence of hybrid breakdown in crosses between *D. subquinaria* and *D. recens* indicates that such strong (lethal) incompatibilities are not at work in the focal study system.

To summarize, from the known characteristics of the *D. subquinaria*/*D. recens* system, no significant stabilizing effect on the infection polymorphism is to be expected.

6.3.2 Exemplary equilibrium states without hybrid male sterility

In this section, we exclude the two background loci, A and B, from the general model described in Section 6.2 and Appendix C. Thus, we can gauge the effect that the infection polymorphism alone has on the spread of female mating preference. We exemplify that in the four-population stepping-stone model, dynamics occur that in principal are equivalent to those that we have found in the last chapter for two parapatric host populations.

Fisherian runaways

Very much like in the two-population scenario of the last chapter, runaway processes take place if mating preference is not costly. Furthermore, the type of runaway can be determined by our "rule of thumb" (see Eq. (5.32)): If the rejection probability of the mating preference allele, P_1, is larger than the coefficient of viability selection at the trait locus, i.e., if $p_1 > s$ then the preferred trait becomes fixed while the preference allele spreads to the same frequency in all four populations. For the opposite relation, $p_1 < s$, the preference allele spreads to fixation in all populations. In Tab. 6.2, we present exemplary equilibrium allele frequencies for both cases. Parameters are the same for both examples, except that in Tab. 6.2a, $p_1 = 1$, whereas in Tab. 6.2b, $p_1 = 0.05$.

In Section 5.3.1, we compared the numbers of generations it takes P_1 to get from 10% to 90% of its equilibrium frequency as a measure for the rate of spread of female mating preference. This method was designed to be relatively independent of the introduction frequency of the mutant allele (0.1%) and the arbitrary precision by which we define an equilibrium to be reached (10^{-7}). In order to qualitatively compare different scenarios in this chapter, we equate this generation difference to the time it takes the female mating preference to spread. For convenience, we also use the term "rate of spread" in a synonymous way. In this sense, we can state that the two exemplary runaway processes of Tab. 6.2 take (a) 2704 and (b) 83685 generations, respectively. The first type of runaway process that results in

(a) Fixation of the preferred trait

	Population			
Allele	1	2	3	4
T_1	0.	0.	0.	0.
T_2	1.	1.	1.	1.
T_3	0.	0.	0.	0.
T_4	0.	0.	0.	0.
P_0	0.817	0.817	0.817	0.817
P_1	0.183	0.183	0.183	0.183
U	0.999	0.956	0.022	0.02
W	0.001	0.044	0.978	0.98

(b) Fixation of the preference allele

	Population			
Allele	1	2	3	4
T_1	0.935	0.038	0.001	0.
T_2	0.065	0.951	0.041	0.003
T_3	0.	0.011	0.91	0.046
T_4	0.	0.001	0.048	0.952
P_0	0.	0.	0.	0.
P_1	1.	1.	1.	1.
U	0.999	0.965	0.022	0.02
W	0.001	0.035	0.978	0.98

Table 6.2: Fisherian runaways. The table gives equilibrium allele frequencies for two exemplary sets of parameter values that result in the spread of P_1 (row highlighted) and runaway selection. In the upper table (a), the trait preferred by P_1 females becomes fixed in all four populations (row highlighted). This type of runaway occurs if $p_1 > s$ (compare Eq. (5.32)). In the bottom table (b), the preference allele reaches fixation in all populations. This occurs if $p_1 < s$. The spread of the P_1 allele from 10% to 90% of its equilibrium frequency (compare *rates of spread* in Section 5.3.1) takes 2704 and 83685 generations, respectively. Parameters: (a) $p_1 = 1$; (b) $p_1 = 0.05$; (both) $f = 0, l_{\mathrm{CI}} = 0.9, m = 0.005, r = 0.5, q = 1, s = 0.1, t = 0.87$.

the fixation of the preferred trait is therefore considerably faster than the second type which yields the fixation of the preference allele in all populations. This is due to the stronger mating preference which yields stronger sexual selection pressures on the trait T_2.

In neither of these runaway processes, premating isolation is established. In particular, the pattern of reproductive character displacement that is observed in the *D. subquinaria*/*D. recens* system does not evolve.

Reinforcement of premating isolation

Our model can produce such a pattern if mating preferences incur costs. In general, very slight costs suffice to inhibit runaway processes and limit the spread of the mating preference mutant allele, P_1, to population 2. If the trait and preference loci fully recombine, however, the preference allele never spreads to high frequencies. For one exemplary set of parameter values, $f = 0$, $l_{\mathrm{CI}} = 0.9$, $m = 0.005$, $p_1 = 1$, $q = 0.99$, $r = 0.5$, $s = 0.1$, $t = 0.87$, Tab. 6.3a shows equilibrium allele frequencies in all four populations. In population 2, P_1 reaches 12.2%, whereas it does not spread in neither of the other populations. Thus, premating isolation is reinforced albeit only to a minor degree.

Reinforcement is significantly enhanced by reduced recombination between the trait and preference loci, as can be seen in Tab. 6.3b, but runaway processes also occur more easily. Premating isolation becomes stronger because the equilibrium frequency of P_1 in population 2 is higher; also, spread happens faster. With the same parameter values as above but $r = 0$ and $q = 0.95$, P_1 spreads to 90% in population 2 instead of 12%, and the spread takes merely 1888 instead of 18234 generations.

Stronger selection at the trait locus favors reinforcement. For instance, with still the same set of parameter values but with $s = 1$ instead of $s = 0.1$, P_1 spreads to an equilibrium frequency of 21% in population 2 if the T and P loci fully recombine ($r = 0.5$), and to fixation (equilibrium frequency of 99.7%) if they are tightly linked ($r = 0$).

Low levels of recombination and strong trait selection pressures also allow the female mating preference to spread despite considerable costs, i.e., low transition rates, q. For $r = 0.5$ and $s = 0.1$, spread of P_1 is impeded altogether if $q \leq 0.95$. But for $r = 0$ and $s = 1$, P_1 still spreads to fixation in population 2 and reinforces premating isolation if $q = 0.7$

On a side note, in comparison with Chapter 5, the more complex structure of four stepping-stone populations adds a new qualitative behavior of the system as the transition from Fisherian runaway to reinforcement is made on decreasing q. For instance, consider the following set of parameter values: $f = 0$, $l_{\mathrm{CI}} = 0.9$, $m = 0.005$, $r = 0$, $p_1 = 1$, $s = 0.1$, $t = 0.87$. Then, if $0.97 \leq q \leq 1$, a runaway process proceeds through all four populations which drives T_1 to fixation in all of them. Reinforcement occurs if $0.92 \leq q \leq 0.95$, and if $q < 0.92$ then P_1 goes to extinction in all populations. But for the range $0.95 < q < 0.97$, a new kind of runaway process takes place that is restricted to the two uninfected populations, driving the preferred trait to fixation and P_1 to the same frequency in both of them whereas the two infected populations are unaffected. However, as this example illustrates, this type of runaway process is restricted to very small regions of the parameter space.

Our simulations show that if for a set of parameter values mating preference spreads then usually arbitrarily weak mating preference is sufficient for the spread to take place. However, the weaker the preference is, the more slowly does the spread proceed. We also established by further extending our model to allow for more alleles at the preference locus that mutant alleles with low rejection probabilities become replaced by alleles with stronger preference. This means that a succession of mutations is possible that results in a stepwise increase of mating preferences. In

(a) Full recombination

	Population			
Allele	1	2	3	4
T_1	0.922	0.027	0.001	0.
T_2	0.078	0.963	0.043	0.002
T_3	0.	0.009	0.909	0.045
T_4	0.	0.	0.048	0.952
P_0	0.993	0.878	0.997	1.
P_1	0.007	0.122	0.003	0.
U	0.999	0.965	0.022	0.02
W	0.001	0.035	0.978	0.98

(b) No recombination

	Population			
Allele	1	2	3	4
T_1	0.934	0.005	0.	0.
T_2	0.066	0.991	0.015	0.
T_3	0.	0.004	0.937	0.047
T_4	0.	0.	0.048	0.953
P_0	0.983	0.099	0.992	1.
P_1	0.017	0.901	0.008	0.
U	1.	0.967	0.022	0.02
W	0.	0.033	0.978	0.98

Table 6.3: Reinforcement by costly mating preference. The tables give equilibrium allele frequencies for two exemplary sets of parameter values for which P_1 spreads only in population 2 (highlighted cells). This yields some degree of premating isolation. P_1 spreads faster and reaches higher frequencies if recombination between trait and preference loci is reduced. The spread of the P_1 allele from 10% to 90% of its equilibrium frequency (compare *rates of spread* in Section 5.3.1) takes 18234 and 1888 generations, respectively. Parameters: (a) $r = 0.5$, $q = 0.99$; (b) $r = 0$, $q = 0.95$; (both) $f = 0$, $l_{\mathrm{CI}} = 0.9$, $m = 0.005$, $p_1 = 1$, $s = 0.1$, $t = 0.87$.

the long (evolutionary) run, only two types of females prevail, non-discriminating ones and those with the strongest preference for T_2 males.

6.3.3 Reinforcement in the full model

In the full model, i.e., including HMS resulting from recombinant nucleotypes at the A and B background loci, the findings of the previous Section 6.3.2 in principal hold. In Tab. 6.4, the equilibrium allele frequencies of the system are presented for the same set of parameter values as in Tab. 6.3b, extended by complete HMS, $h = 1$. The preference allele, P_1, spreads to a frequency of 90.7% in population 2 whereas

	Population			
Allele	1	2	3	4
T_1	0.928	0.005	0.	0.
T_2	0.072	0.991	0.007	0.
T_3	0.	0.004	0.945	0.047
T_4	0.	0.	0.048	0.953
P_0	0.981	0.093	0.995	1.
P_1	0.019	0.907	0.005	0.
$A_1/B_1{}^*$	1.	0.995	0.008	0.
$A_2/B_2{}^*$	0.	0.005	0.992	1.
U	1.	0.968	0.021	0.02
W	0.	0.032	0.979	0.98

Table 6.4: Reinforcement in the full model. The table gives equilibrium allele frequencies for the same exemplary set of parameter values as in Tab. 6.3b, with $h = 1$ added. Divergence at the trait and background loci is stable, as is the infection polymorphism. P_1 spreads only in population 2 (highlighted cell). The spread of the P_1 allele from 10% to 90% of its equilibrium frequency takes 1506 generations (compare *rates of spread* in Section 5.3.1). *Background loci A and B are interchangeable. Parameters: $f = 0$, $h = 1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.005$, $r = 0$, $p_1 = 1$, $q = 0.95$, $s = 0.1$, $t = 0.87$.

no spread occurs in either of the other populations. In Fig. 6.5, the prevalent nucleocytotypes within each population can be seen. Overall, the separation of genotypes and cytotypes is remarkably pronounced: The infection polymorphism is stable. Divergence between all four populations at the trait locus is maintained, as is divergence at the two background loci between populations 1 and 2 on the one hand and populations 3 and 4 on the other. Thus, premating isolation is reinforced, and the pattern of reproductive character displacement within *D. subquinaria* observed in the real-world system is reproduced to some degree.

The reinforcement process occurs at a slightly faster rate than without HMS: Comparing Tables 6.4 and 6.3b, the P_1 allele spreads in 1506 and 1888 generations, respectively. This is due to the fact that females in population 2 that harbor the P_1 allele not only benefit from avoiding CI matings but also have the advantage that they produce less hybrid offspring and thus less sterile sons than wildtype females with the P_0 allele. In fact, an important difference to all previous scenarios is that female preference for the trait T_2 would spread in the infected population 3. This is because in contrast to cytoplasmic incompatibility, hybrid male sterility in our model is symmetric.

Gene flow reduction. In order to numerically measure effective migration rates and gene flow factors in the four-population model, we extended the method from Section 4.2.5 to be applicable to metapopulations that consist of more than two populations (see Appendix D). The pattern of gene flow reduction in the stepping-stone model before and after reinforcement is depicted in Fig. 6.5. It should be kept in mind that immigrant inviability effects due to divergent selection at the trait locus

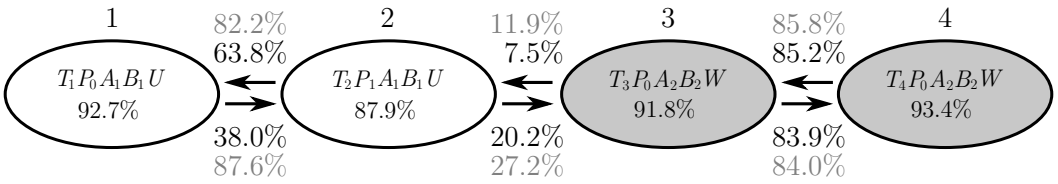


Figure 6.5: Most frequent nucleocytotypes and gene flow reduction. In reference to the model scenario (see Fig. 6.3), this figure shows for each of the four populations the respective predominant nucleocytotype (and its frequency in percent) for the equilibrium state presented in Tab. 6.4. Gene flow factors (in percent) for each migratory connection between the populations are given for the equilibrium state before (values in gray) and after reinforcement (black). Parameters: $f = 0$, $h = 1$, $l_{\mathrm{CI}} = 0.9$, $m = 0.005$, $r = 0$, $p_1 = 1$, $q = 0.95$, $s = 0.1$, $t = 0.87$.

reduce gene flow for migration between all populations (compare Section 5.3.4). The extent of this reduction can be seen in the gene flow factors for migration between populations 3 and 4 and also in the factors for migration between populations 1 and 2 before reinforcement. Only the reduction surpassing these approximately 15% can be attributed to cytoplasmic incompatibility and sexual selection.

Gene flow between the infected populations is largely unaffected by the spread of P_1. However, both gene flow between the uninfected populations and through the hybrid zone is reduced in both directions of migration. This is because both gene flow into and out of the "choosy" population 2 is impeded. Interestingly, effective migration out of population 2 into population 1 is more than halved, i.e., within-species gene flow is very strongly affected. In conclusion, cytoplasmic incompatibility, hybrid male sterility, and female mating preference effectively combine to reduce gene flow between the species ranges.

6.4 Discussion

In this chapter, we extended our two-population model of CI-driven reinforcement to accommodate the geographic setting of the two mycophagous *Drosophila* species *D. subquinaria* and *D. recens* in North America. Each species was represented by one allopatric and one parapatric population, yielding a stepping-stone model with four populations. We estimated parameter values from the literature to adjust our model. Thus, we set up the conditions under which reinforcement would have to occur if CI was indeed the underlying cause of reproductive character displacement in the focal system.

We found that the hybrid male sterility apparent in crossings of the two species does not stabilize the *Wolbachia* infection polymorphism. In real species, HMS is often caused by incompatibilities between more than two gene loci (see Table 8.2 in Coyne and Orr, 2004). This could amplify synergistic effects between HMS and CI somewhat, but we do not expect them to significantly increase critical migration rates.

Furthermore, we showed that females in the uninfected parapatric population benefit from preferring locally adapted males because they have better adapted offspring, avoid CI matings, and have less sterile sons. Thus, patterns of sexual selection similar to those observed in the real system can evolve due to the spread

of a female mating preference allele, P_1. Premating isolation both between species and within the uninfected species is reinforced. For reinforcement to be possible, female mating preferences must incur (slight but not too large) costs. Reinforcement is favored by strong viability selection at the trait locus that functions as a sexual cue. Most importantly, high equilibrium frequencies of P_1 are only attained if the gene loci for the trait and the female preference are in close physical linkage and thus recombination is strongly reduced. If conditions are favorable, the rate at which female mating preference is established ($\sim$ 2000 generations for the set of parameter values in Tab.6.4) conforms with the empirical time frame we estimated to be about 90 000 generations.

Compared to the two-population scenario of Chapter 5, the parameter regions for reinforcement are much smaller in the four population stepping-stone scenario. Indeed, there are stringent conditions that must be met for reinforcement to occur. However, this is consistent with previous findings (see, e.g., Liou and Price, 1994). In general, a problem arises in spatial versions of reinforcement models, since selection pressure for prezygotic isolation may be restricted to a hybrid zone, and gene flow from allopatric populations reduce the probability of reinforcement. In effect, migration of wildtype females from allopatric populations constitutes a selective force acting strongly against the spread of mating preference in the parapatric population. This is however not restricted to *Wolbachia*-induced CI but rather a general problem of reinforcement.

In the focal system, uninfected *D. subquinaria* females that are sympatric with infected *D. recens* show much stronger mate discrimination than allopatric ones. In fact, they not only entirely reject heterospecific males but also strongly avoid matings with conspecific but allopatric males. Our model does not include a means for females to distinguish between species. However, such "species recognition" could be implemented by an additional female mating preference, e.g., by an allele for assortative mating with males that share the female's allele composition at the two background loci. In concordance with the studies of Coyne and Orr (1989) and Coyne and Orr (1997) who found that pre- and postzygotic reproductive isolation between allopatric *Drosophila* species evolve at similar rates, such mate discrimination could be the result of nuclear divergence in allopatry and not of reinforcement. Taken together, such species recognition and female trait preference could produce exactly the patterns observed in *D. subquinaria* and *D. recens*.

Our model is one of haploid sexual organisms. To what extent can our results explain processes in real, i.e., diploid hosts of *Wolbachia*? Hilgenböcker (2009) has combined in diploid population genetic models effects of both nuclear and cytoplasmic incompatibilities. In these models, synergistic effects can result in mutual stabilization of both nuclear and cytoplasmic divergence between host populations. However, the nuclear incompatibilities (NI) must result in strong hybrid lethality for such effects to occur. If, in concordance with Haldane's rule, only one sex is affected by the incompatibilities, effective stabilization only occurs in taxa where females (i.e., the *Wolbachia* transmitting sex) is heterogametic. This is in line with the results reported in this chapter. In the case of *D. subquinaria* and *D. recens*, hybrid males are completely sterile (but no lethality is observed), whereas hybrid females are fully viable and fertile, so that synergistic effects are not incurred.

Our model assumption of infinitely large and perfectly mixing (except for sex-

ual selection) populations might seem overly unrealistic. However, already Sewall Wright showed theoretically for continuously distributed organisms with low dispersal ranges that "if the effective size [of a population] is 1000, there is only slight differentiation at enormous distances. If it is as large as 10,000 the situation is substantially the same as if there were panmixia throughout any conceivable range." (Wright, 1943, page 124) On the empirical side, using Wright's (1978) F-statistics, Shoemaker and Jaenike (1997) found very little within species genetic differentiation for *D. recens* and another mycophagous quinaria species, *D. falleni*, in Northeast America. This suggests high rates of gene flow and essentially panmictic populations despite the patchiness of resources that the flies rely on. Furthermore, for a greenhouse population of *D. limbata* which is a non-mycophagous member of the quinaria species complex, Hummel et al. (1979) estimated effective sizes in the order of tens of thousands of flies. They extrapolated from their data that the total maximum population size in the local area was three million. Together, these studies are consistent with our assumption of infinitely large and panmictic populations of *D. recens* and *D. subquinaria*.

The salient weakness of our model then is the seemingly weak migratory link between these large panmictic populations. We argue, however, that migration rates in the order of a few per mill or even per cent should not be considered weak migration if populations are very large. This is simply because the absolute number of migrants is $N_e m$ so that small migration rates and large effective population sizes yield considerable amounts of migration. Indeed, our simulations show that for two populations linked by migration at rates of 0.001 in both directions and without isolation barrier, equilibrium allele frequencies at a weakly selected locus are attained within a few dozen generations no matter how large the deviation from the equilibrium at the beginning.

The empirical data of the *D. subquinaria*/*D. recens* system is absolutely conclusive in spotting *Wolbachia* as the major determinant of reproductive character displacement. In some sense, we feel that although our model shows that reinforcement is possible, our findings conflict with this conclusiveness. Of course, one possibility is that just the required stringent conditions are met in the focal system. However, we deem it more likely that some feature of the system that stabilizes the infection polymorphism has been overlooked so far. A more stable CI reproductive barrier would allow for greater migration rates which in turn would foster reinforcement by the following logic. With higher rates of migration, infection frequencies in "uninfected" populations would increase (without *Wolbachia* spreading), thus greatly increasing the risk of incompatible matings for uninfected females in the hybrid zone. Consequently, the benefit of avoiding CI matings would be much larger, and mating preferences might spread more easily, i.e., under a broader range of parameter conditions, to higher equilibrium frequencies, and with greater rapidity.

One way that the *Wolbachia* infection polymorphism could be stabilized would be by additional reproductive barriers between *D. subquinaria* and *D. recens* that have synergistic effects on critical migration rates. We suggest that one (or both) of the following extrinsic postzygotic barriers might be involved in *D. subquinaria* and *D. recens* reproductive isolation (compare Tab. 2.4).

(i) If *D. subquinaria* and *D. recens* fill (slightly) different niches then recombinant hybrids might not fit properly in either of these niches. Therefore, hybrids could be

ecologically less viable than pure species individuals (see type III.A(i) in Tab. 2.4). Importantly, such ecological inviability could easily be overlooked in the laboratory as it might only show under natural conditions where the appropriate niches come into play.

(ii) Behavioral sterility would result if recombinant hybrids are less attractive to pure parent species individuals (see type III.A(ii) in Tab. 2.4). This form of sexual selection might be tested for by experiments similar to those conducted by Jaenike et al. (2006) for matings between pure species individuals. Mass population or no-choice trials with females of either *D. subquinaria* or *D. recens* and hybrid males could reveal sexual selection against hybrids.

On the theoretical side, it would be straightforward to extend our model to include both of these presumable isolation barriers. Ecological hybrid inviability could be modeled with the background loci A and B that are already included in the model of this chapter. The pure species allele combinations A_1B_1 and A_2B_2 would then have higher (adult) viability than recombinant A_1B_2 and A_2B_1 hybrids. To model behavioral sterility, female mating preference could be introduced that use the background loci as mating cues, similar to a model by Proulx and Servedio (2009). This would resemble a kind of species recognition and in combination with the processes modeled in the present chapter could produce just the pattern of mating preferences empirically found by Jaenike et al. (2006) (compare Fig. 6.2).

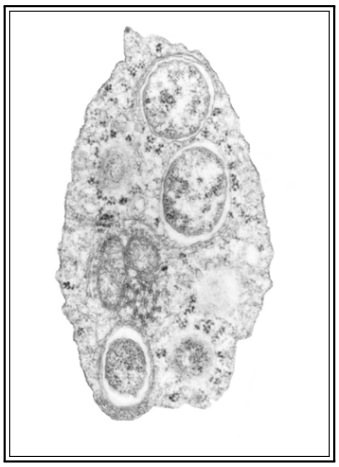
Another possible cause for increased critical migration rates of *Wolbachia* spread (and related to the previous one) could lie in fecundity costs of infection that only accrue in a *D. subquinaria* genetic background. This could be experimentally checked by measuring fertility of infected *D. subquinaria* females produced by transinfection or by hybridization with *D. recens* and repeated backcrossing (also see below).

The most intriguing possible cause of increased stability of the infection polymorphism, however, comes from a follow-up study to Jaenike et al. (2006). Jaenike (2007) introgressed *Wolbachia* from *D. recens* into *D. subquinaria* by hybridization and repeated backcrossing of female offspring with *D. subquinaria* males. This lead to the emergence of a male-killing (MK) phenotype in allopatric but not in sympatric populations of *D. subquinaria*. While the author focuses on the host genetic influences on MK expression, it is striking that the CI phenotype that *Wolbachia* has in its natural host, *D. recens*, seems to be absent in *D. subquinaria*. In hybrid crossings between *D. recens* males and *D. subquinaria* females, however, cytoplasmic incompatibility is fully expressed, suggesting that *D. subquinaria* females are not capable of rescuing the sperm modification implemented by *Wolbachia*. The failure to cause CI when introgressed into *D. subquinaria* might thus be explained by the presence of a nuclear suppressor of *Wolbachia*'s *mod* function in *D. subquinaria* males. Koehncke et al. (2009) have shown in a theoretical model that male-specific modifier alleles that lower *Wolbachia*'s ability to modify sperm (e.g., by decreasing *Wolbachia* density in testes) rapidly spread and can result in the population-wide loss of a *Wolbachia* infection. Diversity of mtDNA haplotypes within *D. subquinaria* suggests that this species has not been infected with *Wolbachia* in the recent past (Shoemaker et al., 1999) which argues against such suppressor genes. On the other hand, it is undisputed that host genetic background can influence CI expression (e.g., Poinsot et al., 1998; Sinkins et al., 2005). Therefore, the approach taken by Koehncke et al. (2009) could still be adopted to model suppression of CI modifica-

tion in *D. subquinaria* males.

On a cautionary note, without being checked against (e.g., by metabolic costs) such modifier alleles should also spread in *D. recens* populations which would drive the *Wolbachia* infection to extinction in the study system altogether. Preliminary numerical simulations of a model including CI suppression with costs that are only incurred by males that actually harbor a *Wolbachia* infection show that an increase of the critical migration rate by a factor of ten is possible, e.g., from $m_{\text{crit}} \approx 0.005$ to $m_{\text{crit}} \approx 0.05$. This significantly increases the benefit of avoiding CI matings, and in fact our simulations show that mating preference in the parapatric *D. subquinaria* population can reach practical fixation in this case. Further simulations are necessary, however, to confirm this scenario and to analyze whether reinforcement is attained in a broader region of the parameter space. In all cases, we strongly suggest that empirical studies be undertaken that investigate the possible existence of CI suppression in *D. subquinaria*.

In summary, we have demonstrated in this chapter theoretically that *Wolbachia*-induced unidirectional CI could have promoted reinforcement of female mating preferences in the *D. subquinaria* and *D. recens* system. However, to produce the patterns of sexual selection observed in laboratory mating experiments, rather specific conditions have to be met. We have raised concerns that some important features of the system might have been overlooked so far, and we have outlined several empirical as well as theoretical approaches to further elucidate this fascinating speciation process.



Chapter 7

Summary and perspectives

The central objective of this thesis was to study the role of *Wolbachia*-induced unidirectional cytoplasmic incompatibility (CI) in host speciation processes. We reviewed the literature on the biological background relevant for our analysis in Chapter 2. Our approach included three cases of mathematical modeling to analyze different aspects of *Wolbachia*'s possible role (Chapters 3 to 5) and one case of applying the models to a real-world speciation process (Chapter 6).

In Chapter 3, we analyzed the conditions that enable stable infection polymorphisms, i.e., coexistence of uninfected populations next to populations infected with *Wolbachia*. For mainland-island scenarios, we showed analytically that if the maternal transmission of *Wolbachia* is imperfect or if infected females suffer fecundity costs then a stable infection polymorphism is possible provided that migration is below a critical rate. We demonstrated that these critical migration rates can serve as estimates of the stability of infection patterns in the case of two-way migration. For symmetric migration, the maximal stability is attained if levels of CI are intermediate. We discussed special cases of the general model in which either *Wolbachia* transmission was perfect, infected females did not incur fecundity costs, or cytoplasmic incompatibility was complete.

Chapter 4 was concerned with the impact that a stable infection polymorphism has on inter-population gene flow. We developed fitness graphs for the reproductive values of infected and uninfected host organisms. From the reproductive values, we analytically calculated effective migration rates for mainland-island scenarios as a measure of gene flow reduction due to unidirectional CI. Utilizing numerical simulations, we corroborated that these effective rates can be used to estimate gene flow in more complex population structures. Gene flow is reduced asymmetrically; in the extreme case of complete cytoplasmic incompatibility, it is reduced to zero in the direction from infected to uninfected populations and to 50% in the opposite direction. This asymmetry turns infected populations into population genetic sinks and uninfected populations into sources. Therefore, in larger population structures, uninfected populations can maintain local adaptations in the face of migration con-

siderably better than infected populations.

In Chapter 5, we combined models of *Wolbachia* dynamics with simple models of host population genetics to investigate whether unidirectional CI selects for mating preferences that allow females to avoid incompatible matings if an infection polymorphism between populations exists. We showed that in uninfected populations that receive migration from an infected mainland, premating isolation readily evolves following the spread of a mutant allele at a locus for female mating preference. We derived a heuristic "effective selection coefficient" to explain the spread of the allele and showed that a minimum strength of mating preference is required that depends on the level of CI. A major finding for scenarios with two-way migration was that costs of female mating preference are crucial. If costs are too high, evolution of mating preference is barred, but if they are too small, runaway selection of mating preference and preferred trait erases nuclear differences between populations. We analyzed the conditions that favor reinforcement of premating isolation and discussed its effect on gene flow.

In Chapter 6, we applied our reinforcement model to the real-world speciation of *Wolbachia*-infected *Drosophila recens* and uninfected *D. subquinaria* in North America which show patterns of sexual selection that fit the expected outcome of reinforcement due to unidirectional CI. We found that our model can produce the observed patterns but only if stringent conditions are satisfied. In particular, female mating preferences must incur costs, and recombination between trait and preference loci must be strongly reduced. Based on these findings, we suggested that important features of the *Drosophila* system might have been overlooked so far that stabilize the *Wolbachia* infection polymorphism and thus have favored reinforcement.

There are several tie-in topics which lend themselves to fruitful continuation of our research. First of all, our models are fully deterministic, any stochasticity is completely neglected. In general, most model parameters (e.g., CI levels, migration rates, and transmission rates) could and in natural systems will vary stochastically. Taking a stochastic modeling approach, Reuter et al. (2008) argue that population sub-division facilitates invasions of *Wolbachia* infections. For bidirectional CI, Branca et al. (2009) showed that stochasticity generally decreases the stability of infection polymorphisms but that it might in fact be increased if divergent viability selection at a trait locus is strong. It would be interesting to determine the effects of stochasticity on the reinforcement processes outlined in this thesis.

A second shortcoming of our models is that it remains uncertain how well our findings translate from haploid to diploid organisms. Although results on the stability of infection polymorphisms look promising, it is impossible to analyze in our models for instance how dominance effects in mating preferences alter reinforcement dynamics. Diploid models therefore would be natural extensions of our approach.

In the discussion of the previous chapter (see Section 6.4), we have already suggested some interesting possibilities of future research on the *D. subquinaria*/*D. recens* system. As argued there, our model indicates that some feature of this system might have been overlooked so far. This feature would presumably greatly stabilize the *Wolbachia* infection polymorphism. Possible candidates include ecological

viability costs of hybrids and —most intriguingly— the presence of CI suppressor genes in *D. subquinaria*. Theoretical modeling of these possibilities should not prove too difficult and could provide important insights. However, research of these issues would certainly benefit from close collaboration with experimental groups. In fact, further investigation might also provide insight into the molecular basis of CI that remains unknown to date despite considerable efforts.

Are there any other study systems to further our knowledge of *Wolbachia*'s role in reinforcement processes? At first sight, the *Drosophila testacea* species group presents a good candidate for such a system (compare Fig. 6.1). However, there is no geographic overlap between infected and uninfected species as *D. neotestacea* and *D. putrida* occur in North America, *D. testacea* is found throughout continental Eurasia, and *D. orientacea* is endemic to Japan. In fact, crossing experiments with the two uninfected species *D. testacea* and *D. neotestacea* and the infected species *D. orientacea* revealed different forms of reproductive isolation, including sexual selection, failure to transfer sperm, and hybrid inviability; cytoplasmic incompatibility, however, seemed to play no role in isolation (Grimaldi et al., 1992). In general, these findings are in concordance with the conclusions of Coyne and Orr (1989, 1997). These authors screened over 171 pairs of *Drosophila* species and concluded that sym- or parapatric pairs show larger mate discrimination than allopatric pairs, but that neither pre- nor postzygotic isolation evolves consistently faster in allopatric populations of *Drosophila*. Therefore, we suggest a re-analysis of these data with infection status added. Such a comparative study may reveal whether earlier evolution of mate discrimination is associated with closely related species pairs where one is infected with *Wolbachia*. A whole new set of study systems for the interaction of cytoplasmic incompatibility and premating isolation might be exposed and provide a much more solid empirical basis for theoretical modeling efforts than is currently available.

In closing, we hope to have convincingly argued for *Wolbachia*'s role in speciation processes of their hosts. These fascinating bacteria might not be the only or even major players in arthropod speciation but for their (evolutionary) life time achievements, they may well be honored with an *Academy Award for Best Actors in a Supporting Role.*

Appendices

A Fixpoints of *Wolbachia* dynamics with local host adaptation

For the sake of completeness, we document here the fixpoints of the infection dynamics of *Wolbachia* in a mainland-island scenario where local selection acts on a host trait locus (see Section 3.2.6). For both scenarios with an uninfected or infected mainland, the critical migration rates can be determined by setting $x_2^* = x_3^*$ and solving for m (see Equations (3.53) and (3.54)). The formulas are typeset without much change to the original output by *Wolfram's Mathematica* (Wolfram, 2008), i.e., only minor efforts have been made to rearrange them for better readability.

A.1 Uninfected mainland

In a host population that receives migration from an uninfected mainland where natural selection favors a different trait, the equilibrium *Wolbachia* frequencies are

$$x_1^* = 0, \tag{A.1a}$$

$$\begin{aligned} x_{2,3}^* = \Big[& l_{\mathrm{CI}}+f+(l_{\mathrm{CI}}+f)(1-m)s \mp \Big((1+(1-m)s)\cdot \\ & \big((l_{\mathrm{CI}}-f)^2(1+(1-m)s) - 4l_{\mathrm{CI}}(1-f)(1+s-ms- \\ & f(1-m)(1+s))t - 4l_{\mathrm{CI}}(1-f)^2(1-m)(1+s)t^2\big)\Big)^{\frac{1}{2}}\Big] \Big/ \\ & \big(2l_{\mathrm{CI}}(1-m)(1+s)(f(1-t)+t)\big). \end{aligned} \tag{A.1b}$$

A.2 Infected mainland

If the mainland is infected, the equilibrium infection frequencies in the island population are

$$x_1^* = x_{\mathrm{main}}, \quad \text{as per } x_3^* \text{ in Eq. (3.2b)}, \tag{A.2a}$$

$$\begin{aligned} x_{2,3}^* = \Big[& (1-f)(1-m)(1+s)\big(l_{\mathrm{CI}}(1-m)(1+s)- \\ & f(1-m)(1+s) - \sqrt{R}\,(1+m+s-ms)\big)t \pm \\ & \Big(2\,(1-f)^2(1-m)^2(1+s)^2t^2 \cdot \Big(l_{\mathrm{CI}}^2(-2ms(1+s)+ \\ & (1+s)^2 + m^2\left(1+s^2\right)\Big) - f\Big((1-m)\sqrt{R}\cdot \\ & (1-m(1-s)+s)(1+s) + f\big(-2ms(1+s)+(1+s)^2+ \\ & m^2\left(1+s^2\right)\big)\Big) - l_{\mathrm{CI}}\big((1-m)\sqrt{R}(1+s)(1+m+(1-m)s)+ \\ & 2(1+m+s-ms)^2t + 2f^2(1-m)^2(1+s)^2(1-t)t- \\ & 2(1-m)^2(1+s)^2t^2 + 4f\big(m(1+s-ms) - \big(-2ms(1+s)+ \end{aligned} \tag{A.2b}$$

$$(1+s)^2+m^2\left(1+s^2\right))t+(1-m)^2(1+s)^2t^2))\Big)^{\frac{1}{2}}\Big]\Big/$$
$$\Big(4l_{\mathrm{CI}}(1-f)(1-m)^2(1+s)^2(f(1-t)+t)t\Big),$$

where

$$R=R(f,l_{\mathrm{CI}},t)=(l_{\mathrm{CI}}-f)^2-4l_{\mathrm{CI}}(1-f)^2(1-t) \tag{A.3}$$

is the function introduced in Eq. (3.3).

B Reproductive values in a heterogeneous host population

In a heterogeneous host population with infected and uninfected *Wolbachia* hosts coexisting, the reproductive values (ν) of uninfected females (subscript f,u) and males (m,u), and infected females (f,w) and males (m,w) (see Section 4.2.3) are very cumbersome. For completeness, however, they are shown in this appendix section throughout which the function R, given by Eq. (A.3), will be used.

B.1 Uninfected hosts

For the uninfected hosts in the heterogeneous population, the reproductive values of females ($\nu_{\mathrm{f,u}}$) and males ($\nu_{\mathrm{m,u}}$) are:

$$\nu_{\mathrm{f,u}}=\left[2(-1+l_{\mathrm{CI}})l_{\mathrm{CI}}(-1+f)t\right]\Big/ \tag{B.4a}$$
$$\Big[f\left(f+\sqrt{R}\right)+l_{\mathrm{CI}}^2(1+(-1+f)t)-$$
$$l_{\mathrm{CI}}\Big(\sqrt{R}+f^2(5-6t)t+\left(4-3\sqrt{R}-6t\right)t+$$
$$f\left(2+3t\left(-3+\sqrt{R}+4t\right)\right)\Big)\Big],$$
$$\nu_{\mathrm{m,u}}=l_{\mathrm{CI}}\Big[l_{\mathrm{CI}}^2+(-1+f)\left(f+\sqrt{R}\right)(1+2f(-1+t)-2t)t+ \tag{B.4b}$$
$$l_{\mathrm{CI}}\left(-f+\sqrt{R}-3t+(7-4f)ft+4(-1+f)^2t^2\right)\Big]\Big/$$
$$\Big[l_{\mathrm{CI}}^3+f^2\left(f+\sqrt{R}\right)+l_{\mathrm{CI}}^2(-f+\sqrt{R}-4t+(9-5f)ft+$$
$$5(-1+f)^2t^2)+l_{\mathrm{CI}}(5f^3(-1+t)t+\sqrt{R}\,t(-2+3t)+ft$$
$$\left(-4+5\sqrt{R}+5t-6\sqrt{R}\,t\right)+$$
$$f^2\left(-1+t\left(9-3\sqrt{R}-10t+3\sqrt{R}\,t\right)\right))\Big].$$

The average reproductive value of an uninfected host (resident or immigrant) then calculates to

$$\nu_{\mathrm{u}}=\frac{1}{2}\left(\nu_{\mathrm{f,u}}+\nu_{\mathrm{m,u}}\right)$$
$$=(-1+f)t\Big(l_{\mathrm{CI}}^2+f\left(f-\sqrt{R}\right)- \tag{B.5}$$

$$l_{\mathrm{CI}}\left(\sqrt{R}+4t+2\left(f+f(-5+3f)t-3(-1+f)^2t^2\right)\right)\Big/$$
$$2\Big((l_{\mathrm{CI}}-f)^2+2(-1+f)\left(l_{\mathrm{CI}}(2+l_{\mathrm{CI}})-4l_{\mathrm{CI}}f+f^2\right)t-$$
$$3l_{\mathrm{CI}}(-1+f)^2(-4+3f)t^2+9l_{\mathrm{CI}}(-1+f)^3t^3\Big).$$

B.2 Infected hosts

The reproductive values of the infected members of heterogeneous population are very cumbersome formulas and read —for females ($\nu_{\mathrm{f,w}}$) and males ($\nu_{\mathrm{m,w}}$), respectively:

$$\begin{aligned}\nu_{\mathrm{f,w}} = \; & -4l_{\mathrm{CI}}\Bigg[(l_{\mathrm{CI}}^4(1+2f(-1+t)-2t)+(-1+f)f\left(f+\sqrt{R}\right)\cdot \\ & (1+2f(-1+t)-2t)\left(f^2(-1+t)-t\right)t+l_{\mathrm{CI}}^3(\sqrt{R}- \\ & 2\left(3+\sqrt{R}\right)t+13f^3(-1+t)^2t+17t^2-13t^3+f^2(4+ \\ & t\left(-34+69t-39t^2\right))+f(-2\left(1+\sqrt{R}\right)+ \\ & t\left(27+2\sqrt{R}-60t+39t^2\right)))+l_{\mathrm{CI}}(-1+f)t\cdot \\ & \Big(10f^5(-1+t)^3t+f^4(-1+t)^2(-1+3t(9-2\sqrt{R}+ \\ & 2\left(-6+\sqrt{R}\right)t))+t\Big(-3\sqrt{R}+t(2+9\sqrt{R}+t(-6- \\ & 11\sqrt{R}+4t+6\sqrt{R}t))\Big)-ft(2-11\sqrt{R}+t(1+39\sqrt{R}+ \\ & 3t\left(-3-17\sqrt{R}+2t+8\sqrt{R}t\right)))-f^3(-1+t)(\sqrt{R}+ \\ & t(4\left(-7+4\sqrt{R}\right)+t\left(67-41\sqrt{R}-44t+24\sqrt{R}t\right)))+ \\ & f^2\Big(-\sqrt{R}+t(11-17\sqrt{R}+t(-35+69\sqrt{R}+t(39- \\ & 87\sqrt{R}+4\left(-4+9\sqrt{R}\right)t)))\Big)+l_{\mathrm{CI}}^2\Big(20f^5(-1+t)^3t^2- \\ & f^4(-1+t)^2t(8+t(-81+100t))+t(-4\sqrt{R}+t(9+11\sqrt{R}- \\ & t\left(31+9\sqrt{R}+t(-41+20t)\right)))+f^3(-1+t)(2+ \\ & t\left(-9\left(2+\sqrt{R}\right)+t\left(157+9\sqrt{R}+4t(-81+50t)\right)\right))+ \\ & f^2\Big(1+2\sqrt{R}-t(16+22\sqrt{R}+t(-148-47\sqrt{R}+ \\ & t\left(423+27\sqrt{R}-486t+200t^2\right)))\Big)+f\Big(-\sqrt{R}+t(4+17\sqrt{R}+ \\ & t\left(-59-40\sqrt{R}+t\left(3\left(61+9\sqrt{R}\right)+4t(-56+25t)\right)\right)\right)\Big)\Bigg]\Big/ \\ & \Bigg[\left(l_{\mathrm{CI}}+f+\sqrt{R}\right)(l_{\mathrm{CI}}-f+\sqrt{R}-2t+2(3-2f)ft+\end{aligned} \tag{B.6a}$$

$$
\begin{aligned}
&4(-1+f)^2t^2)\Big(l_{\mathrm{CI}}^3+f^2\left(f+\sqrt{R}\right)+l_{\mathrm{CI}}^2(-f+\sqrt{R}-4t+\\
&(9-5f)ft+5(-1+f)^2t^2)+l_{\mathrm{CI}}(5f^3(-1+t)t+\\
&\sqrt{R}t(-2+3t)+ft\left(-4+5\sqrt{R}+5t-6\sqrt{R}t\right)+\\
&f^2\left(-1+t\left(9-3\sqrt{R}-10t+3\sqrt{R}t\right)\right)\Big)\Big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
\nu_{\mathrm{m,w}} = 4l_{\mathrm{CI}}\Big[&l_{\mathrm{CI}}^4+(-1+f)f\left(f+\sqrt{R}\right)(1+2f(-1+t)-2t)t\cdot \qquad \text{(B.6b)}\\
&\Big(f^2+(-1+(4-3f)f)t+2(-1+f)^2t^2\Big)+l_{\mathrm{CI}}^3(-2f+\\
&\sqrt{R}-6t-f(-13+f(6+f))t+(-1+f)^2(7+2f)t^2-\\
&(-1+f)^3t^3)+l_{\mathrm{CI}}^2\Big[-4f^5(-1+t)^3t^2+f^4(-1+t)^2t^2(3+20t)-\\
&f^3(-1+t)t\left(6-\sqrt{R}+t\left(-41+\sqrt{R}+4t(3+10t)\right)\right)+\\
&t\left(-4\sqrt{R}+t\left(9+5\sqrt{R}+t\left(-23+\sqrt{R}+t(11+4t)\right)\right)\right)-\\
&f\Big(\sqrt{R}+t(-4-9\sqrt{R}+t(47+8\sqrt{R}+t(-91+3\sqrt{R}+\\
&4t(8+5t)))))\Big)+f^2\Big(1+t(-2\left(5+2\sqrt{R}\right)+t(78+\sqrt{R}+\\
&t\left(3\left(-41+\sqrt{R}\right)+2t(9+20t)\right)))\Big)\Big]+l_{\mathrm{CI}}(-1+f)t\cdot\\
&\Big[2f^5(-1+t)^3t(-3+4t)-t(-1+2t)(-3\sqrt{R}+t(2+\\
&3\sqrt{R}+t\left(-6+\sqrt{R}+4t\right)))-f^4(-1+t)^2(-1+t(11-\\
&2\sqrt{R}+2t\left(-28+\sqrt{R}+20t\right)))+f^3(-1+t)\Big(\sqrt{R}+\\
&t\left(-4\left(1+\sqrt{R}\right)+t\left(83-5\sqrt{R}+4t\left(-41+2\sqrt{R}+20t\right)\right)\right)\Big)+\\
&ft\Big(-2+13\sqrt{R}+t(-11-29\sqrt{R}+t(67+9\sqrt{R}+2t(-47+\\
&4\sqrt{R}+20t)))\Big)+f^2\Big(\sqrt{R}+t(5-17\sqrt{R}+t(43+25\sqrt{R}+\\
&t\left(3\left(-61+\sqrt{R}\right)-4t\left(-54+3\sqrt{R}+20t\right)\right)))\Big)\Big]\Big]\Big/\\
&\Big[\Big(l_{\mathrm{CI}}+f+\sqrt{R}\Big)(l_{\mathrm{CI}}-f+\sqrt{R}-2t+2(3-2f)ft+\\
&4(-1+f)^2t^2)\Big(l_{\mathrm{CI}}^3+f^2\left(f+\sqrt{R}\right)+l_{\mathrm{CI}}^2(-f+\sqrt{R}-4t+\\
&(9-5f)ft+5(-1+f)^2t^2)+l_{\mathrm{CI}}(5f^3(-1+t)t+\\
&\sqrt{R}t(-2+3t)+ft\left(-4+5\sqrt{R}+5t-6\sqrt{R}t\right)+\\
&f^2\left(-1+t\left(9-3\sqrt{R}-10t+3\sqrt{R}t\right)\right)\Big)\Big)\Big].
\end{aligned}
$$

The average reproductive value of an infected host therefore is

$$
\begin{aligned}
\nu_{\mathrm{w}} &= \frac{1}{2}\left(\nu_{\mathrm{f,w}} + \nu_{\mathrm{m,w}}\right) \\
&= 4l_{\mathrm{CI}}(f(-1+t)-t)\Bigg[-l_{\mathrm{CI}}^4 + (-1+f)f\left(f+\sqrt{R}\right)\cdot \\
&\quad (1+2f(-1+t)-2t)(f(-1+t)-t)t - l_{\mathrm{CI}}^3(-2f+ \\
&\quad \sqrt{R}-5t+(12-7f)ft+7(-1+f)^2t^2) - l_{\mathrm{CI}}(-1+f)t\cdot \\
&\quad \Big[-4f^4(-2+t)(-1+t)^2t + t^2(-2+3\sqrt{R}-2t(-3+ \\
&\quad 2\sqrt{R}+2t)) + f^3(-1+t)(-1+t(19-4\sqrt{R}+2t\cdot \\
&\quad \left(-19+2\sqrt{R}+8t\right))) + f(\sqrt{R}+(-1+t)t(3-2\sqrt{R}+ \\
&\quad 2t\left(-9+6\sqrt{R}+8t\right))) - f^2\Big(\sqrt{R}+t(-15+6\sqrt{R}+ \\
&\quad t\left(56-19\sqrt{R}+6t\left(-11+2\sqrt{R}+4t\right)\right)\Big)\Big)\Big] - l_{\mathrm{CI}}^2\Big(12f^4\cdot \\
&\quad (-1+t)^2t^2 + f^3t\left(4+t\left(-43+87t-48t^2\right)\right) + t(-3\sqrt{R}+ \\
&\quad t\left(4+5\sqrt{R}+3t(-5+4t)\right)) - f(\sqrt{R}+t(-2-8\sqrt{R}+ \\
&\quad t\left(27+10\sqrt{R}-69t+48t^2\right))) + f^2(1+t(-6-5\sqrt{R}+ \\
&\quad t\left(54+5\sqrt{R}+9t(-13+8t)\right)))\Big)\Bigg] \Bigg/ \\
&\quad \Bigg[\left(l_{\mathrm{CI}}+f+\sqrt{R}\right)\left(l_{\mathrm{CI}}-f+\sqrt{R}-2t+2(3-2f)ft+\right. \\
&\quad 4(-1+f)^2t^2\Big)\Big(l_{\mathrm{CI}}^3 + f^2\left(f+\sqrt{R}\right) + l_{\mathrm{CI}}^2(-f+\sqrt{R}-4t+ \\
&\quad (9-5f)ft+5(-1+f)^2t^2) + l_{\mathrm{CI}}(5f^3(-1+t)t+ \\
&\quad \sqrt{R}t(-2+3t) + ft\left(-4+5\sqrt{R}+5t-6\sqrt{R}t\right) + \\
&\quad \left.f^2\left(-1+t\left(9-3\sqrt{R}-10t+3\sqrt{R}t\right)\right)\right)\Big)\Bigg].
\end{aligned}
\tag{B.7}
$$

C Mathematical *Drosophila* model

Basically, we use the difference equations developed in Chapter 5 to describe the life cycle of *Drosophila* flies in Chapter 6. For more explanations on parameter definitions, see the model section in Chapter 5.

For subsequent generations, let $x_{i,j,k,l,m,n}$ and $x'_{i,j,k,l,m,n}$ denote the frequencies of nucleocytotype $T_jP_kA_lB_mC_n$ in population i. According to the life cycle, the frequencies $x'_{i,j,k,l,m,n}$ can be calculated in four steps. Intermediate frequencies are $x^{+}_{i,j,k,l,m,n}$ after migration, and $x^{++}_{i,j,k,l,m,n}$ after viability selection (because it is mathematically more convenient, sexual selection is included in the reproduction step).

C.1 Migration

In a stepping stone structure (see Fig. 6.3), migration occurs only between neighboring populations. We assume that a fraction m of a population is replaced by migrants from each population that it receives migration from:

$$x^{+}_{1,\vec{\gamma}} = (1-m)\, x_{1,\vec{\gamma}} + m\, x_{2,\vec{\gamma}}, \tag{C.1a}$$

$$x^{+}_{2,\vec{\gamma}} = m\, x_{1,\vec{\gamma}} + (1-2\,m)\, x_{2,\vec{\gamma}} + m\, x_{3,\vec{\gamma}}, \tag{C.1b}$$

$$x^{+}_{3,\vec{\gamma}} = m\, x_{2,\vec{\gamma}} + (1-2\,m)\, x_{3,\vec{\gamma}} + m_{3,4}\, x_{4,\vec{\gamma}}, \tag{C.1c}$$

$$x^{+}_{4,\vec{\gamma}} = m\, x_{3,j,k,l} + (1-m)\, x_{4,\vec{\gamma}}, \tag{C.1d}$$

where the short notation $\vec{\gamma} = (j,k,l,m,n)$ has been used. $x^{+}_{i,\vec{\gamma}} = x^{+}_{i,j,k,l,m,n}$ are the frequencies of nucleocytotype $T_jP_kA_lB_mC_n$ in population i after migration.

C.2 Viability selection

In each population i, the trait T_i is favored by local viability selection before all other traits (compare Eq. (5.2)):

$$x^{++}_{i,j,k,l,m,n} = \begin{cases} \dfrac{1+s}{W_i}\; x^{+}_{i,j,k,l,m,n} & \text{if } i=j \\ \dfrac{1}{W_i}\; x^{+}_{i,j,k,l,m,n} & \text{else,} \end{cases} \tag{C.2}$$

where W_i is the sum over indexes j, k, l, m, and n of all numerators in Eq. C.2 for population i. $x^{++}_{i,j,k,l,m,n}$ are the frequencies of nucleocytotype $T_jP_kA_lB_mC_n$ in population i after viability selection, and s is called the selection coefficient.

C.3 Sexual selection

The weight for sexual selection defined in Eq. (5.6) has to be slightly modified in order to account for P_1 females preferring T_2 males (instead of T_1) in the four population model since T_2 is the locally adaptive trait in the sympatric uninfected population:

$$S_{i,j,k} = \begin{cases} 1 & \text{if } j=0 \\ \dfrac{1}{1-p_j\, q\left(1-y^{++}_{i,j}\right)} & \text{if } j=1 \text{ and } k=2 \\ \dfrac{1-p_j}{1-p_j\, q\left(1-y^{++}_{i,j}\right)} & \text{if } j=1 \text{ and } k\neq 2. \end{cases} \tag{C.3}$$

Here $y^{++}_{i,j}$ denotes the frequency of trait T_j after viability selection in population i, and $S_{i,j,k}$ yields the relative number of matings between P_j-females and T_k-males in that population. The parameters p_i and q are the rejection probability of P_i and the transition probability, respectively.

C.4 Reproduction

The reproduction step takes into account sexual selection, female fecundity, *Wolbachia* transmission, cytoplasmic incompatibility, nuclear inheritance, and hybrid

male sterility. The recurrence equation for reproduction is (compare Eq. (5.14)):

$$x'_{i,j,k,l,m,n} = \frac{1}{W_i} \sum_{\alpha,\zeta} \sum_{\beta,\eta} \sum_{\gamma,\vartheta} \sum_{\delta,\iota} \sum_{\varepsilon,\kappa} x^{++}_{i,\alpha,\beta,\gamma,\delta,\varepsilon}(\text{♀})\, x^{++}_{i,\zeta,\eta,\vartheta,\iota,\kappa}(\text{♂}) \tag{C.4a}$$
$$\cdot\, S_{i,\beta,\zeta}\, F_\varepsilon\, L_{\kappa,n}\, V_{\varepsilon,n}\, H_{\vartheta,\iota}\, I^{TP}_{\alpha,\beta,\zeta,\eta,j,k}\, I^{AB}_{\gamma,\delta,\vartheta,\iota,l,m}$$
$$W_i = \sum_{j,k,l,m,n} \sum_{\alpha,\zeta} \sum_{\beta,\eta} \sum_{\gamma,\vartheta} \sum_{\delta,\iota} \sum_{\varepsilon,\kappa} x^{++}_{i,\alpha,\beta,\gamma,\delta,\varepsilon}(\text{♀})\, x^{++}_{i,\zeta,\eta,\vartheta,\iota,\kappa}(\text{♂}) \tag{C.4b}$$
$$\cdot\, S_{i,\beta,\zeta}\, F_\varepsilon\, L_{\kappa,n}\, V_{\varepsilon,n}\, H_{\vartheta,\iota}\, I^{TP}_{\alpha,\beta,\zeta,\eta,j,k}\, I^{AB}_{\gamma,\delta,\vartheta,\iota,l,m}.$$

Equations C.4 give the frequencies of $T_j P_k A_l B_m C_n$ offspring that is produced from matings between females of nucleocytotype $T_\alpha P_\beta A_\gamma B_\delta C_\varepsilon$ and males of nucleocytotype $T_\zeta P_\eta A_\vartheta B_\iota C_\kappa$ in population i. We have marked the frequencies of females (♀) and males (♂) for clarity; the double plus superscript implies that female and male frequencies after viability selection have to be used.

Most weights in Equations C.4 can be adopted unchanged from Chapter 5:

- F_i as per Eq. (5.10): ***female fecundity***
- $V_{i,j}$ as per Eq. (5.11): ***cytotype inheritance*** (or: transmission of *Wolbachia*)
- $L_{i,j}$ as per Eq. (5.12): ***cytoplasmic incompatibility***
- $I^{TP}_{i,j,k,l,m,n} = I_{i,j,k,l,m,n}$ as per Eq. (5.13): ***nuclear inheritance at the trait and preference loci***, we use the superscript TP to denote the loci.

For ***nuclear inheritance at the background loci***, A and B, the same weight $I^{AB}_{i,j,k,l,m,n} = I_{i,j,k,l,m,n}$ as per Eq. (5.13) is used, the recombination rate in this case is assumed to be $r = 0.5$. Here, the female genotype is $(A_i B_j)$, the male genotype is $(A_k B_l)$, and the offspring's genotype is $(A_m B_n)$.

The one weight that is new, $H_{i,j}$, reflects ***hybrid male sterility***. In what follows, we describe its rationale in more detail. In principal, the weight is a modified version of a well-studied model for two-locus incompatibilities (Servedio, 2000). To be precise, two background loci, A and B, interact epistatically to cause low sterility in recombinant males. There are two allele combinations, A_1B_1 and A_2B_2, that have evolved by divergence in allopatry from an ancestral combination A_0B_0. In males, the reciprocal recombinants, A_1B_2 and A_2B_1, do not have viability effects but cause complete sterility. Females with these genotypes, however, are fully viable and fertile. We use running indices to denote these alleles: A_i, where $i \in \{1, 2\}$, are the alleles at the background nuclear locus A, and B_i, where again $i \in \{1, 2\}$, are the alleles at the background nuclear locus B. Then, hybrid male sterility can be expressed by a weighting factor,

$$H_{i,j} = \begin{cases} 1 - h & \text{if } i \neq j \\ 1 & \text{else.} \end{cases} \tag{C.5}$$

Here, i and j mark the male nuclear background allele combination, A_iB_j, in a male mating partner. The parameter h is the level of sterility and could in principal range between zero and one. Throughout Chapter 6, we assume that $h = 1$ to meet the fact that the sterility in the focal study system is complete.

If we were to model sex-independent hybrid inviability instead, we could use the same weight but in the recurrence equation for the reproduction step, Eq. (C.4), the indices would indicate the offspring's background alleles, i.e., we would use $H_{l,m}$ instead of $H_{\vartheta,\iota}$.

The mathematical model described in this appendix section is far too complex to be amenable to analytical solutions, so that we resorted to numerical simulations. In contrast to Chapter 5, however, the computing time proved to be too long for rigorous screenings of the parameter space. We therefore restricted the analysis to certain sets of parameter values and starting conditions (see Sections 6.2 and 6.3).

D Effective migration rates in large metapopulations

In Section 4.2.5, we described a numerical method to calculate effective migration rates for two populations linked by migration. Here, we present an algorithm that extends the method to allow for the measurement of gene flow in structures consisting of more than two populations.

1. Start with a metapopulation at equilibrium.
2. Introduce a neutral marker locus with three alleles, N_1, N_2, and N_3.
3. Choose two populations i and j that effective migration rates are to be calculated for. These are termed "focal" populations, whereas all other populations are "background" populations. In one focal population, N_1 is fixed at the beginning, in the other focal population, N_2 is fixed. In all background populations, N_3 is fixed.
4. After each migration step, any allele at the N locus in a background population is turned into N_3. By contrast, in the two focal populations any fraction that an N_3-nucleocytotype takes is distributed to the corresponding N_1- and N_2-nucleocytotypes such that their ratio is unchanged (the precise mathematical formulation of this redistribution process is given below in Equations (D.1)). Otherwise, the metapopulation's generation cycle remains unchanged.
5. Let this system run for one hundred generations, and from the focal populations, insert the N_1 allele frequencies of generations 99 and 100 into Equations (4.34). This yields $m_{ij,\mathrm{eff}}$ for effective migration from population j into i and $m_{ji,\mathrm{eff}}$ for the opposite direction.
6. Repeat steps 2 to 5 for each pair of populations in the metapopulation that is actually connected by migration.

Redistribution at the neutral marker locus in the focal populations. For each nucleocytotype $T_jP_kA_lB_mC_n$ of the model without the neutral locus (see Appendix C), three neutral extensions with either of the three neutral alleles exist in a focal population after migration: $N_1T_jP_kA_lB_mC_n$, $N_2T_jP_kA_lB_mC_n$, and $N_3T_jP_kA_lB_mC_n$. Denoting their frequencies by $n_{i,\vec{\gamma}}$, where $\vec{\gamma} = (j,k,l,m,n)$, we can formulate the

redistribution as follows:

$$n_{1,\vec{\gamma}}^{(\text{redist.})} = n_{1,\vec{\gamma}} + n_{3,\vec{\gamma}} \frac{n_{1,\vec{\gamma}}}{n_{1,\vec{\gamma}} + n_{2,\vec{\gamma}}} \tag{D.1a}$$

$$n_{2,\vec{\gamma}}^{(\text{redist.})} = n_{2,\vec{\gamma}} + n_{3,\vec{\gamma}} \frac{n_{1,\vec{\gamma}}}{n_{1,\vec{\gamma}} + n_{2,\vec{\gamma}}} \tag{D.1b}$$

For example, consider the following frequencies after migration in one of the focal populations: $n_{1,\vec{\gamma}} = 0.06$, $n_{2,\vec{\gamma}} = 0.03$, and $n_{3,\vec{\gamma}} = 0.01$. Then, the redistribution yields $n_{1,\vec{\gamma}}^{(\text{redist.})} = 0.0\bar{6}$ and $n_{2,\vec{\gamma}}^{(\text{redist.})} = 0.0\bar{3}$.

Acknowledgements

You would not be reading this thesis if it were not for quite a couple of people who supported me. First and foremost, I want to thank Peter Hammerstein. Apart from giving me the chance to pursue my research at the Institute for Theoretical Biology (ITB), he was the one who introduced me to the fascinating field of theoretical biology. I will always remember his excellent lectures on evolutionary game theory, and after all it was a particularly nice one of his hand drawn overhead transparencies that sparked my interest in *Wolbachia*. In our numerous discussions, I have learned to especially value his skill of drawing my attention away from petty detail toward key issues.

Numerous were also the fruitful and productive, *Wolbachia*-related and unrelated discussions I had with Arndt Telschow who also commented on earlier drafts of this thesis. It was always contagious to see Jack Werren nearly burst with enthusiasm for the funny little *Wolbachia* bacteria. These two have contributed invaluable ideas and feedback to my research.

During my time at the ITB, I have had the pleasure to work with a lot of people, and by work I mean not only actually working on joint projects, but also discussing old and new scientific ideas, or just sitting in the same office and building a wall of (emptied) espresso boxes. This applies in no particular order to Samuel Glauser, Tim Oppermann, Edward Hagen, Nicole Hess, Benjamin Bossan, Roman Zug, Simone Monreal, Arndt Telschow, Kirsten Hilgenböcker, Sofia Figureido, Arnulf Köhncke, Judith Dannowski, and Victor Anaya. Because I always appreciated after-work card playing distractions, I want to thank the core "Doppelkopf" players at the ITB, Kirsten Hilgenböcker, Robert Schmidt, and Andreas Hantschmann, for letting me participate from time to time.

I would like to thank my parents and my sister Katrin, and special thanks go to my brother Stefan for his proof-reading and baby-sitting services.

Last, but certainly not least, I am especially grateful to Saskia and Boyke for being there.

Bibliography

S. G. E. Andersson, A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. M. Alsmark, R. M. Podowski, A. K. Naslund, A. Eriksson, H. H. Winkler, and C. G. Kurland. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature*, 396:133–140, 1998. 1, 9

L. Baldo, S. Bordenstein, J. J. Wernegreen, and J. H. Werren. Widespread recombination throughout *Wolbachia* genomes. *Molecular Biology and Evolution*, 23: 437–449, 2006a. 9

L. Baldo, J. C. D. Hotopp, K. A. Jolley, S. R. Bordenstein, S. A. Biber, R. R. Choudhury, C. Hayashi, M. C. J. Maiden, H. Tettelin, and J. H. Werren. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Applied and Environmental Microbiology*, 72:7098–7110, 2006b. 8, 9

N. Barton and B. O. Bengtsson. The barrier to genetic exchange between hybridising populations. *Heredity*, 57:357–376, 1986. 64

C. B. Beard, S. L. O'Neill, R. B. Tesh, F. F. Richards, and S. Aksoy. Modification of arthropod vector competence via symbiotic bacteria. *Parasitology Today*, 9: 179–183, 1993. 8, 62

B. O. Bengtsson. The flow of genes through a genetic barrier. In J. J. Greenwood, H. Harvey, and M. Slatkin, editors, *Evolution: essays in honour of John Maynard Smith*, pages 31–42. Cambridge University Press, Cambridge, 1985. 64, 66

S. R. Bordenstein. Bordenstein Laboratory homepage, 2008. URL `http://bordensteinlab.vanderbilt.edu`. 6

S. R. Bordenstein. Symbiosis and the origin of species. In K. Bourtzis and T. A. Miller, editors, *Insect symbiosis*, pages 283–304. CRC Press, Boca Raton, 2003. 22, 84, 86

S. R. Bordenstein, F. P. O'Hara, and J. H. Werren. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature*, 409:707–710, 2001. 86

B. Bossan. *The first insect cell cycle,* Wolbachia*, and the mechanism of cytoplasmic incompatibility*. Diploma thesis, Humboldt University, Berlin, Germany, 2009. 15

D. Bouchon, T. Rigaud, and P. Juchault. Evidence for widespread *Wolbachia* infection in isopod crustaceans: molecular identification and host feminization. *Proceedings of the Royal Society London B*, 265:1081–1090, 1998. 17

K. Bourtzis, A. Nirgianaki, G. Markakis, and C. Savakis. *Wolbachia* infection and cytoplasmic incompatibility in *Drosophila* species. *Genetics*, 144:1063–1073, 1996. 29

K. Bourtzis, H. R. Braig, and T. L. Karr. Cytoplasmic incompatibility. In K. Bourtzis and T. A. Miller, editors, *Insect Symbiosis*, pages 217–246. CRC Press, Boca Raton, 2003. 8, 16

L. Boyle, S. L. O'Neill, H. M. Robertson, and T. L. Karr. Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*. *Science*, 260:1796–1799, 1993. 11

H. R. Braig, W. Zhou, S. L. Dobson, and S. L. O'Neill. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *Journal of Bacteriology*, 180:2373–2378, 1998. 11

A. Branca, F. Vavre, J. Silvain, and S. Dupas. Maintenance of adaptive differentiation by *Wolbachia*-induced bidirectional cytoplasmic incompatibility: the importance of sib-mating and genetic systems. *BMC Evolutionary Biology*, 9: 185, 2009. 142

J. A. J. Breeuwer and G. Jacobs. *Wolbachia*: intracellular manipulators of mite reproduction. *Experimental and Applied Acarology*, 20:421–434, 1996. 24

J. A. J. Breeuwer and J. H. Werren. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature*, 346: 558–560, 1990. 11, 24

M. Bulmer. Structural instability of models of sexual selection. *Theoretical Population Biology*, 35:195–206, 1989. 21, 118

G. Callaini and M. G. Riparbelli. Fertilization in *Drosophila melanogaster*: centrosome inheritance and organization of the first mitotic spindle. *Developmental Biology*, 176:199–208, 1996. 15

G. Callaini, R. Dallai, and M. G. Riparbelli. *Wolbachia*-induced delay of paternal chromatin condensation does not prevent maternal chromosomes from entering anaphase in incompatible crosses of *Drosophila simulans*. *Journal of Cell Science*, 110:271–280, 1997. 16

E. Caspari and G. S. Watson. On the evolutionary importance of cytoplasmic sterility in mosquitoes. *Evolution*, 13:568–570, 1959. 2, 8, 28, 61

S. Charlat, J. Engelstädter, E. A. Dyson, E. A. Hornett, A. Duplouy, P. Tortosa, N. Davies, G. K. Roderick, N. Wedell, and G. D. D. Hurst. Competing selfish genetic elements in the butterfly *Hypolimnas bolina*. *Current Biology*, 16:2453–2458, 2006. 24

J. C. Choe and B. J. Crespi. *The evolution of mating systems in insects and arachnids*. Cambridge University Press, Cambridge, 1997. 118

D. J. Clancy and A. A. Hoffmann. Behavior of *Wolbachia* endosymbionts from *Drosophila simulans* in *Drosophila serrata*, a novel host. *American Naturalist*, 149:975–988, 1997. 11

M. E. Clark, Z. Veneti, K. Bourtzis, and T. L. Karr. The distribution and proliferation of the intracellular bacteria *Wolbachia* during spermatogenesis in *Drosophila*. *Mechanisms of Development*, 111:3–15, 2002. 11

M. E. Clark, Z. Veneti, K. Bourtzis, and T. L. Karr. *Wolbachia* distribution and cytoplasmic incompatibility during sperm development: the cyst as the basic cellular unit of CI expression. *Mechanisms of Development*, 120:185–198, 2003. 16

J. A. Coyne and H. A. Orr. The evolutionary genetics of speciation. *Philosophical Transactions of the Royal Society B*, 353:287–305, 1998. 22

J. A. Coyne and H. A. Orr. Patterns of speciation in *Drosophila*. *Evolution*, 43: 362–381, 1989. 137, 143

J. A. Coyne and H. A. Orr. "Patterns of speciation in *Drosophila*" revisited. *Evolution*, 51:295–303, 1997. 22, 137, 143

J. A. Coyne and H. A. Orr. *Speciation*. Sinauer Associates, Inc., Sunderland, Massachusetts, 2004. 18, 19, 20, 22, 61, 86, 127, 136

C. F. Curtis. Population replacement in *Culex fatigans* by means of cytoplasmic incompatibility: 2. Field cage experiments with overlapping generations. *Bulletin of the World Health Organization*, 53:107–119, 1976. 8

C. F. Curtis and T. Adak. Population replacement in *Culex fatigans* by means of cytoplasmic incompatibility: 1. Laboratory experiments with non-overlapping generations. *Bulletin of the World Health Organization*, 51:249–255, 1974. 8

H. da Rocha Lima. Zur Aetiologie des Fleckfiebers. *Berliner klinische Wochenschrift*, 53:567–569, 1916. 6

C. R. Darwin. *The descent of man, and selection in relation to sex*. John Murray, London, 2nd edition, 1874. 20

C. R. Darwin. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Murray, London, 1859. 18

F. Dedeine, F. Vavre, F. Fleury, B. Loppin, M. E. Hochberg, and M. Boulétreau. Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp. *Proceedings of the National Academy of Sciences of the USA*, 98: 6247–6252, 2001. 11

F. Dedeine, M. Bouletreau, and F. Vavre. *Wolbachia* requirement for oogenesis: occurrence within the genus *Asobara* (Hymenoptera, Braconidae) and evidence for intraspecific variation in *A. tabida*. *Heredity*, 95:394–400, 2005. 11

S. L. Dobson, K. Bourtzis, H. R. Braig, B. F. Jones, W. Zhou, F. Rousset, and S. L. O'Neill. *Wolbachia* infections are distributed throughout insect somatic and germ line tissues. *Insect Biochemistry and Molecular Biology*, 29:153–160, 1999. 10

S. L. Dobson, E. J. Marsland, and W. Rattanadechakul. Mutualistic *Wolbachia* infection in *Aedes albopictus*: accelerating cytoplasmic drive. *Genetics*, 160:1087–1094, 2002a. 11

S. L. Dobson, E. J. Marsland, Z. Veneti, K. Bourtzis, and S. L. O'Neill. Characterization of *Wolbachia* host cell range via the in vitro establishment of infections. *Applied and Environmental Microbiology*, 68:656–660, 2002b. 10

T. Dobzhansky. A critique of the species concept in biology. *Philosophy of Science*, 2:344–355, 1935. 64

T. Dobzhansky. *Genetics and the origin of species*. Columbia University Press, New York, 1937. 18, 19, 64, 127, 180

T. Dobzhansky. A review of some fundamental concepts and problems of population genetics. *Cold Spring Harbor Symposia on Quantitative Biology*, 20:1–15, 1955. 122

T. Dobzhansky. Speciation as a stage in evolutionary divergence. *American Naturalist*, 74:312, 1940. 2, 21, 87

O. Duron, C. Bernard, S. Unal, A. Berthomieu, C. Berticat, and M. Weill. Tracking factors modulating cytoplasmic incompatibilities in the mosquito *Culex pipiens*. *Molecular Ecology*, 15:3061–3071, 2006a. 24

O. Duron, P. Fort, and M. Weill. Hypervariable prophage WO sequences describe an unexpected high number of *Wolbachia* variants in the mosquito *Culex pipiens*. *Proceedings of the Royal Society London B*, 273:495–502, 2006b. 16

O. Duron, P. Fort, and M. Weill. Influence of aging on cytoplasmic incompatibility, sperm modification and *Wolbachia* density in *Culex pipiens* mosquitoes. *Heredity*, 98:368–374, 2007. 16

E. A. Dyson and G. D. D. Hurst. Persistence of an extreme sex-ratio bias in a natural population. *Proceedings of the National Academy of Sciences of the USA*, 101: 6520–6523, 2004. 13

L. K. Elfring, J. M. Axton, D. D. Fenger, A. W. Page, J. L. Carminati, and T. L. Orr-Weaver. Drosophila PLUTONIUM protein is a specialized cell cycle regulator required at the onset of embryogenesis. *Molecular Biology of the Cell*, 8:583–593, 1997. 12

J. Engelstädter and A. Telschow. Cytoplasmic incompatibility and host population structure. *Heredity*, 103:196–207, 2009. 8, 23

J. Engelstädter, A. Telschow, and P. Hammerstein. Infection dynamics of different *Wolbachia*-types within one host population. *Journal of Theoretical Biology*, 231: 345–355, 2004. 62, 84

T. L. Erwin. Tropical forests: their richness in Coleoptera and other arthropod species. *Coleopterists Bulletin*, 36:74–75, 1982. 22

J. Felsenstein. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, 35:124–138, 1981. 24

P. M. Ferree and W. Sullivan. A genetic test of the role of the maternal pronucleus in *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila melanogaster*. *Genetics*, 173:839–847, 2006. 15

P. M. Ferree, H. M. Frydman, J. M. Li, J. Cao, E. Wieschaus, and W. Sullivan. *Wolbachia* utilizes host microtubules and dynein for anterior localization in the *Drosophila* oocyte. *PLoS Pathogens*, 1:e14, 2005. 11

P. E. M. Fine. On the dynamics of symbiote-dependent cytoplasmic incompatibility in culicine mosquitoes. *Journal of Invertebrate Pathology*, 31:10–18, 1978. 8, 28, 29, 31, 33, 61, 89

R. A. Fisher. *The genetical theory of natural selection*. Clarendon Press, Oxford, 1930. 20, 21, 65, 76, 86, 117, 119, 180

K. D. Floate, G. K. Kyei-Poku, and P. C. Coghlin. Overview and relevance of *Wolbachia* bacteria in biocontrol research. *Biocontrol Science and Technology*, 16:767, 2006. 8

M. Flor, P. Hammerstein, and A. Telschow. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and the stability of infection polymorphism in parapatric host populations. *Journal of Evolutionary Biology*, 20:696–706, 2007. 27, 29, 30, 47

J. Foster, M. Ganatra, I. Kamal, J. Ware, K. Makarova, N. Ivanova, A. Bhattacharyya, V. Kapatral, S. Kumar, J. Posfai, T. Vincze, J. Ingram, L. Moran, A. Lapidus, M. Omelchenko, N. Kyrpides, E. Ghedin, S. Wang, E. Goltsman, V. Joukov, O. Ostrovskaya, K. Tsukerman, M. Mazur, D. Comb, E. Koonin, and B. Slatko. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biology*, 3:e121, 2005. 8, 12

H. M. Frydman, J. M. Li, D. N. Robson, and E. Wieschaus. Somatic stem cell niche tropism in *Wolbachia*. *Nature*, 441:509–512, 2006. 11

S. Gavrilets. *Fitness landscapes and the origin of species*. Monographs in Population Biology. Princeton University Press, Princeton, 2004. 66

S. Gavrilets. Hybrid zones with Dobzhansky-type epistatic selection. *Evolution*, 51: 1027–1035, 1997. 64

S. Ghelelovitch. Genetic determinism of sterility in the cross-breeding of various strains of *Culex autogenicus* Roubaud. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 234:2386–2388, 1952. 6, 14

R. Giordano, J. Jackson, and H. Robertson. The role of *Wolbachia* bacteria in reproductive incompatibilities and hybrid zones of *Diabrotica* beetles and *Gryllus* crickets. *Proceedings of the National Academy of Sciences of the USA*, 94:11439–11444, 1997. 24

J. M. Gleason, J. Jallon, J. Rouault, and M. G. Ritchie. Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D. sechellia*. *Genetics*, 171:1789–1798, 2005. 119

A. Grafen. A theory of Fisher's reproductive value. *Journal of Mathematical Biology*, 53:15–60, 2006. 65

D. Grimaldi, A. C. James, and J. Jaenike. Systematics and modes of reproductive isolation in the holarctic *Drosophila testacea* species group (Diptera: Drosophilidae). *Annals of the Entomological Society of America*, 85:671–685, 1992. 143

T. Guillemaud, N. Pasteur, and F. Rousset. Contrasting levels of variability between cytoplasmic genomes and incompatibility types in the mosquito *Culex pipiens*. *Proceedings of the Royal Society London B*, 264:245–251, 1997. 24

D. W. Hall, M. Kirkpatrick, and B. West. Runaway sexual selection when female preferences are directly selected. *Evolution*, 54:1862–1869, 2000. 21

T. Handorf, N. Christian, O. Ebenhöh, and D. Kahn. An environmental perspective on metabolism. *Journal of Theoretical Biology*, 252:530–537, 2008. 8

D. L. Hartl and A. G. Clark. *Principles of Population Genetics*. Sinauer Associates, Inc., Sunderland, Massachusetts, 4th edition, 2007. 89

M. Hertig. The rickettsia, *Wolbachia pipientis* (gen. et sp.n.) and associated inclusions of the mosquito, *Culex pipiens*. *Parasitology*, 28:453–486, 1936. 6, 10

M. Hertig and S. B. Wolbach. Studies on rickettsia-like micro-organisms in insects. *Journal of Medical Research*, 44:329–374, 1924. 6

K. Hilgenböcker. Wolbachia*'s role in classical speciation theory*. PhD thesis, Humboldt University, Berlin, Germany, 2009. 84, 131, 137

K. Hilgenböcker, P. Hammerstein, P. Schlattmann, A. Telschow, and J. H. Werren. How many species are infected with *Wolbachia*? – a statistical analysis of current data. *FEMS Microbiology Letters*, 281:215–220, 2008. 1, 8, 179

M. Hiroki, Y. Kato, T. Kamito, and K. Miura. Feminization of genetic males by a symbiotic bacterium in a butterfly, *Eurema hecabe* (Lepidoptera: Pieridae). *Naturwissenschaften*, 89:167–170, 2002. 13

M. Hiroki, Y. Ishii, and Y. Kato. Variation in the prevalence of cytoplasmic incompatibility-inducing *Wolbachia* in the butterfly *Eurema hecabe* across the Japanese archipelago. *Evolutionary Ecology Research*, 7:931–942, 2005. 28

A. A. Hoffmann and M. Turelli. Cytoplasmic incompatibility in insects. In S. L. O'Neill, A. A. Hoffmann, and J. H. Werren, editors, *Influential passengers*, pages 42–80. Oxford University Press, Oxford, 1997. 29, 62

A. A. Hoffmann, M. Turelli, and L. G. Harshman. Factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. *Genetics*, 126:933–48, 1990. 8, 11, 28, 29, 33

A. A. Hoffmann, D. J. Clancy, and E. Merton. Cytoplasmic incompatibility in Australian populations of *Drosophila melanogaster*. *Genetics*, 136:993–999, 1994. 24

A. A. Hoffmann, M. Hercus, and H. Dagher. Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. *Genetics*, 148:221–232, 1998. 11

E. A. Hornett, S. Charlat, A. M. R. Duplouy, N. Davies, G. K. Roderick, N. Wedell, and G. D. D. Hurst. Evolution of male-killer suppression in a natural population. *PLoS Biology*, 4:e283, 2006. 13

E. A. Hornett, A. M. R. Duplouy, N. Davies, G. K. Roderick, N. Wedell, G. D. D. Hurst, and S. Charlat. You can't keep a good parasite down: evolution of a male-killer suppressor uncovers cytoplasmic incompatibility. *Evolution*, 62:1258–1263, 2008. 13

S. Hoshizaki and T. Shimada. PCR-based detection of *Wolbachia*, cytoplasmic incompatibility microorganisms, infected in natural populations of *Laodelphax striatellus* (Homoptera: Delphacidae) in central Japan: has the distribution of *Wolbachia* spread recently? *Insect Molecular Biology*, 4:237–243, 1995. 24

C. J. Hoskin, M. Higgie, K. R. McDonald, and C. Moritz. Reinforcement drives rapid allopatric speciation. *Nature*, 437:1353–1356, 2005.

J. C. D. Hotopp, M. Lin, R. Madupu, J. Crabtree, S. V. Angiuoli, J. Eisen, R. Seshadri, Q. Ren, M. Wu, T. R. Utterback, S. Smith, M. Lewis, H. Khouri, C. Zhang, H. Niu, Q. Lin, N. Ohashi, N. Zhi, W. Nelson, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, J. Sundaram, S. C. Daugherty, T. Davidsen, A. S. Durkin, M. Gwinn, D. H. Haft, J. D. Selengut, S. A. Sullivan, N. Zafar, L. Zhou, F. Benahmed, H. Forberger, R. Halpin, S. Mulligan, J. Robinson, O. White, Y. Rikihisa, and H. Tettelin. Comparative genomics of emerging human ehrlichiosis agents. *PLoS Genetics*, 2:e21, 2006. 4, 10

J. C. D. Hotopp, M. E. Clark, D. C. S. G. Oliveira, J. M. Foster, P. Fischer, M. C. M. Torres, J. D. Giebel, N. Kumar, N. Ishmael, S. Wang, J. Ingram, R. V. Nene, J. Shepard, J. Tomkins, S. Richards, D. J. Spiro, E. Ghedin, B. E. Slatko, H. Tettelin, and J. H. Werren. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science*, 317:1753–1756, 2007. 12

M. E. Huigens and R. Stouthamer. Parthenogenesis associated with *Wolbachia*. In K. Bourtzis and T. A. Miller, editors, *Insect symbiosis*, pages 247–266. CRC Press, Boca Raton, 2003. 13, 17

H. K. Hummel, W. Delden, and R. H. Drent. Estimation of some population parameters of *Drosophila limbata* V. Roser in a greenhouse. *Oecologia*, 41:135–143, 1979. 138

G. D. D. Hurst and F. M. Jiggins. Male-killing bacteria in insects: mechanisms, incidence, and implications. *Emerging Infectious Diseases*, 6:329–337, 2000. 13

G. D. D. Hurst, F. M. Jiggins, and M. E. N. Majerus. Inherited microorganisms that selectively kill male hosts: the hidden players of insect evolution? In K. Bourtzis and T. A. Miller, editors, *Insect symbiosis*, pages 177–197. CRC Press, Boca Raton, 2003. 16

L. D. Hurst. The incidences and evolution of cytoplasmic male killers. *Proceedings of the Royal Society London B*, 244:91–99, 1991. 13

Y. Iwasa and A. Pomiankowski. Continual change in mate preferences. *Nature*, 377: 420–422, 1995. 21

J. Jaenike. Mycophagous *Drosophila* and their nematode parasites. *American Naturalist*, 139:893–906, 1992. 129

J. Jaenike. Spontaneous emergence of a new *Wolbachia* phenotype. *Evolution*, 61: 2244–2252, 2007. 13, 122, 139

J. Jaenike, K. A. Dyer, C. Cornish, and M. S. Minhas. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biology*, 4:e325, 2006. 4, 24, 28, 61, 86, 122, 124, 125, 127, 128, 129, 139, 180

J. Jallon and J. R. David. Variation in cuticular hydrocarbons among the eight species of the *Drosophila melanogaster* subgroup. *Evolution*, 41:294–302, 1987. 119

A. C. James and J. W. O. Ballard. Expression of cytoplasmic incompatibility in *Drosophila simulans* and its impact on infection frequencies and distribution of *Wolbachia pipientis*. *Evolution*, 54:1661–1672, 2000. 24

F. M. Jiggins, G. D. D. Hurst, C. E. Dolman, and M. E. N. Majerus. High-prevalence male-killing *Wolbachia* in the butterfly *Acraea encedana*. *Journal of Evolutionary Biology*, 13:495–501, 2000. 13

P. Juchault, M. Frelon, D. Bouchon, and T. Rigaud. New evidence for feminizing bacteria in terrestrial isopods - evolutionary implications. *Comptes Rendus de l'Académie des Sciences - Life Sciences*, 317:225–230, 1994. 13

D. Kageyama and W. Traut. Opposite sex-specific effects of *Wolbachia* and interference with the sex determination of its host *Ostrinia scapulalis*. *Proceedings of the Royal Society London B*, 271:251–258, 2004. 14

G. P. Keller, D. M. Windsor, J. M. Saucedo, and J. H. Werren. Reproductive effects and geographical distributions of two *Wolbachia* strains infecting the Neotropical beetle, *Chelymorpha Alternans* Boh. (Chrysomelidae, Cassidinae). *Molecular Ecology*, 13:2405–2420, 2004. 28, 83

M. Kirkpatrick. Sexual selection and the evolution of female choice. *Evolution*, 36: 1–12, 1982. 21, 92, 118

M. Kirkpatrick and N. H. Barton. The strength of indirect selection on female mating preferences. *Proceedings of the National Academy of Sciences of the USA*, 94:1282–1286, 1997. 107

M. Kirkpatrick and V. Ravigné. Speciation by natural and sexual selection: models and experiments. *American Naturalist*, 159:S22–S35, 2002. 20, 24, 25

M. Kirkpatrick and M. J. Ryan. The evolution of mating preferences and the paradox of the lek. *Nature*, 350:33–38, 1991. 21

M. Kirkpatrick and M. R. Servedio. The reinforcement of mating preferences on an island. *Genetics*, 151:865–884, 1999. 21, 83, 127

L. Klasson, T. Walker, M. Sebaihia, M. J. Sanders, M. A. Quail, A. Lord, S. Sanders, J. Earl, S. L. O'Neill, N. Thomson, S. P. Sinkins, and J. Parkhill. Genome evolution of *Wolbachia* strain *w*Pip from the *Culex pipiens* group. *Molecular Biology and Evolution*, 25:1877–1887, 2008. 8, 11

L. Klasson, J. Westberg, P. Sapountzis, K. Naslund, Y. Lutnaes, A. C. Darby, Z. Veneti, L. Chen, H. R. Braig, R. Garrett, K. Bourtzis, and S. G. E. Andersson. The mosaic genome structure of the *Wolbachia w*Ri strain infecting *Drosophila simulans*. *Proceedings of the National Academy of Sciences of the USA*, 106: 5725–5730, 2009. 8, 11

Y. Kobayashi and A. Telschow. The concept of effective migration rate and its application in speciation models. In V. T. Koven, editor, *Population Genetics Research Progress*, pages 151–175. Nova Biomedical Books, 2008. 64, 65, 66, 80, 180

Y. Kobayashi, P. Hammerstein, and A. Telschow. The neutral effective migration rate in a mainland-island context. *Theoretical Population Biology*, 74:84–92, 2008. 65

A. Koehncke, A. Telschow, J. H. Werren, and P. Hammerstein. Life and death of an influential passenger: *Wolbachia* and the evolution of CI-modifiers by their hosts. *PLoS ONE*, 4:e4425, 2009. 9, 139

H. Kose and T. L. Karr. Organization of *Wolbachia pipientis* in the *Drosophila* fertilized egg and embryo revealed by an anti-*Wolbachia* monoclonal-antibody. *Mechanisms of Development*, 51:275–288, 1995. 15

T. Kunisawa. Branching orders among α-proteobacteria and mitochondria inferred from gene transpositions. *Journal of Theoretical Biology*, 234:1–5, 2005. 9

R. C. Lacy. Ecological and genetic responses to mycophagy in Drosophilidae (Diptera). In Q. Wheeler and M. Blackwell, editors, *Fungus/insect relationships: perspectives in ecology and evolution*, pages 286–301. Columbia University Press, New York, 1984. 122, 129

R. C. Lacy. Structure of genetic variation within and between populations of mycophagous *Drosophila*. *Genetics*, 104:81–94, 1983. 129

R. Lande. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution*, 33:234–251, 1979.

R. Lande. Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences of the USA*, 78:3721–3725, 1981. 21, 118

H. Laven. Crossing experiments with *Culex* strains. *Evolution*, 5:370–375, 1951. 1, 6, 14, 24, 179

H. Laven. Cytoplasmic inheritance in *Culex*. *Nature*, 177:141–142, 1956. 6

H. Laven. Speciation in mosquitoes. *Cold Spring Harbor Symposia on Quantitative Biology*, 24:166–173, 1959. 2, 86, 179

H. Laven. Vererbung durch Kerngene und das Problem der außerkaryotischen Vererbung bei *Culex pipiens*. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*, 88:478–516, 1957. 6, 14

P. H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 33:183–212, 1945. 75

L. W. Liou and T. D. Price. Speciation by reinforcement of premating isolation. *Evolution*, 48:1451–1459, 1994. 83, 137

N. Lo, C. Paraskevopoulos, K. Bourtzis, S. L. O'Neill, J. H. Werren, S. R. Bordenstein, and C. Bandi. Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. *International Journal of Systematic and Evolutionary Microbiology*, 57: 654–657, 2007. 4, 9, 10

C. Louis and L. Nigro. Ultrastructural evidence of *Wolbachia* rickettsiales in *Drosophila simulans* and their relationships with unidirectional cross-incompatibility. *Journal of Invertebrate Pathology*, 54:39–44, 1989. 10

A. A. Markov. Extension of the limit theorems of probability theory to a sum of variables connected in a chain (Reprint in appendix b). In *Ronald A. Howard: Dynamic Probabilistic Systems, Vol. 1: Markov Chains*. John Wiley and Sons, New York, 1971. 93

L. S. Maroja, M. E. Clark, and R. G. Harrison. *Wolbachia* plays no role in the one-way reproductive incompatibility between the hybridizing field crickets *Gryllus firmus* and *G. pennsylvanicus*. *Heredity*, 101:435–444, 2008. 24

J. F. Marshall. *The British mosquitoes*. Trustees of the British Museum, London, 1938. 14

G. Martin, P. Juchault, and J. J. Legrand. Presentation of symbiotic intracytoplasmic microorganism of oniscoid *Armadillidium vulgare* later accompanied by intersexuality or total feminization of genetic males in the lygenic line. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 276:2313, 1973. 13

R. M. May. How many species are there on earth? *Science*, 241:1441–1449, 1988. 22

J. Maynard Smith and E. Szathmáry. *The major transitions in evolution*. Oxford University Press, New York, 1997. 1

E. Mayr. *Animal species and evolution*. Belknap Press of Harvard University Press, Cambridge, Massachusetts, 1963. 24, 83

E. Mayr. The biological meaning of species. *Biological Journal of the Linnean Society*, 1:311–320, 1969. 18

E. Mayr. *The growth of biological thought*. Harvard University Press, Cambridge, Massachusetts, 1982. 18

E. Mayr. *Systematics and the origin of species from the viewpoint of a zoologist*. Columbia University Press, New York, 1942. 18, 20, 64

C. J. McMeniman, R. V. Lane, B. N. Cass, A. W. C. Fong, M. Sidhu, Y. Wang, and S. L. O'Neill. Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti*. *Science*, 323:141–144, 2009. 8

H. Merçot and S. Charlat. *Wolbachia* infections in *Drosophila melanogaster* and *D. simulans*: polymorphism and levels of cytoplasmic incompatibility. *Genetica*, 120:51–59, 2004. 24, 29

H. Merçot, B. Llorente, M. Jacques, A. Atlan, and C. Montchamp-Moreau. Variability within the Seychelles cytoplasmic incompatibility system in *Drosophila simulans*. *Genetics*, 141:1015–1023, 1995. 86

K. Min and S. Benzer. *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent, causing degenerization and early death. *Proceedings of the National Academy of Sciences of the USA*, 94:10792–10796, 1997. 8

T. H. Morgan. *The physical basis of heredity*. J. B. Lippincott Company, Philadelphia, 1919. 4

H. J. Muller. Isolating mechanisms, evolution and temperature. *Biological Symposia*, 6:71–125, 1942. 127

S. Narita, M. Nomura, Y. Kato, and T. Fukatsu. Genetic structure of sibling butterfly species affected by *Wolbachia* infection sweep: evolutionary and biogeographical implications. *Molecular Ecology*, 15:1095–1108, 2006. 24

L. Nigro and T. Prout. Is there selection on RFLP differences in mitochondrial DNA? *Genetics*, 125:551–555, 1990. 28

M. A. F. Noor. Reinforcement and other consequences of sympatry. *Heredity*, 83: 503–508, 1999. 21

P. Nosil, T. H. Vines, and D. J. Funk. Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, 59:705–719, 2005. 84, 115

F. Ødegaard. How many species of arthropods? Erwin's estimate revised. *Biological Journal of the Linnean Society*, 71:583–597, 2000. 22

H. Ogata and J. Claverie. Metagrowth: a new resource for the building of metabolic hypotheses in microbiology. *Nucleic Acids Research*, 33:D321–D324, 2005. 8

S. L. O'Neill and T. L. Karr. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. *Nature*, 348:178–180, 1990. 24

S. L. O'Neill, R. Giordano, A. M. Colbert, T. L. Karr, and H. M. Robertson. 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proceedings of the National Academy of Sciences of the USA*, 89:2699–2702, 1992. 6, 9

S. L. O'Neill, M. M. Pettigrew, S. P. Sinkins, H. R. Braig, T. G. Andreadis, and R. B. Tesh. In vitro cultivation of *Wolbachia pipientis* in an *Aedes albopictus* cell line. *Insect Molecular Biology*, 6:33–39, 1997. 8, 10

S. E. Osborne, Y. S. Leong, S. L. O'Neill, and K. N. Johnson. Variation in antiviral protection mediated by different *Wolbachia* strains in *Drosophila simulans*. *PLoS Pathogens*, 5:e1000656, 2009. 11

F. Pérez and B. E. Granger. IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, 9:21–29, 2007. 4

S. J. Perlman and J. Jaenike. Infection success in novel hosts: an experimental and phylogenetic study of *Drosophila*-parasitic nematodes. *Evolution*, 57:544–557, 2003. 4, 123

S. J. Perlman, G. S. Spicer, D. D. Shoemaker, and J. Jaenike. Associations between mycophagous *Drosophila* and their *Howardula* nematode parasites: a worldwide phylogenetic shuffle. *Molecular Ecology*, 12:237–249, 2003. 123

PLoS Biology Synopsis. Genome sequence of the intracellular bacterium *Wolbachia*. *PLoS Biology*, 2:e76, 2004. 4

D. Poinsot, K. Bourtzis, G. Markakis, C. Savakis, and H. Merçot. *Wolbachia* transfer from *Drosophila melanogaster* into *D. simulans*: host effect and cytoplasmic incompatibility relationships. *Genetics*, 150:227–237, 1998. 11, 139

D. Poinsot, S. Charlat, and H. Merçot. On the mechanism of *Wolbachia*-induced cytoplasmic incompatibility: confronting the models with the facts. *BioEssays*, 25:259–265, 2003. 15, 16

J. R. Powell. Interspecific cytoplasmic gene flow in the absence of nuclear gene flow: evidence from *Drosophila*. *Proceedings of the National Academy of Sciences of the USA*, 80:492–495, 1983. 18

D. C. Presgraves. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics*, 163:955–972, 2003. 20

D. C. Presgraves. A genetic test of the mechanism of *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila*. *Genetics*, 154:771–776, 2000. 16

D. C. Presgraves. Patterns of postzygotic isolation in Lepidoptera. *Evolution*, 56: 1168–1183, 2002. 22

T. Price. Sexual selection and natural selection in bird speciation. *Philosophical Transactions of the Royal Society B*, 353:251–260, 1998. 21

S. R. Proulx and M. R. Servedio. Dissecting selection on female mating preferences during secondary contact. *Evolution*, 63:2031–2046, 2009. 118, 139

J. M. Ranz, C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science*, 300: 1742–1745, 2003. 127

J. L. Rasgon and T. W. Scott. *Wolbachia* and cytoplasmic incompatibility in the California *Culex pipiens* mosquito species complex: parameter estimates and infection dynamics in natural populations. *Genetics*, 165:2029–2038, 2003. 8

J. L. Rasgon, C. E. Gamston, and X. Ren. Survival of *Wolbachia pipientis* in cell-free medium. *Applied and environmental microbiology*, 72:6934–6937, 2006. 10

K. M. Reed and J. H. Werren. Induction of paternal genome loss by the paternal-sex-ratio chromosome and cytoplasmic incompatibility bacteria *Wolbachia*: a comparative study of early embryonic events. *Molecular Reproduction and Development*, 40:408–418, 1995. 15

P. Renesto, N. Crapoulet, H. Ogata, B. L. Scola, G. Vestris, J. Claverie, and D. Raoult. Genome-based design of a cell-free culture medium for *Tropheryma whipplei*. *Lancet*, 362:447–449, 2003. 8

M. Reuter, L. Lehmann, and F. Guillaume. The spread of incompatibility-inducing parasites in sub-divided host populations. *BMC Evolutionary Biology*, 8:134, 2008. 142

K. T. Reynolds, L. J. Thomson, and A. A. Hoffmann. The effects of host age, host nuclear background and temperature on phenotypic effects of the virulent *Wolbachia* strain *popcorn* in *Drosophila melanogaster*. *Genetics*, 164:1027–1034, 2003. 29

M. Riegler and S. L. O'Neill. The genus *Wolbachia*. In M. Dworkin, S. Falkow, E. Rosenberg, K. Schleifer, and E. Stackebrandt, editors, *The prokaryotes*, volume 5 of *A handbook on the biology of bacteria*, pages 547–561. Springer, New York, 3rd edition, 2006. 9, 10

M. Riegler and C. Stauffer. *Wolbachia* infections and superinfections in cytoplasmically incompatible populations of the European cherry fruit fly *Rhagoletis cerasi* (Diptera, Tephritidae). *Molecular Ecology*, 11:2425–2434, 2002. 24, 28

M. Riegler, M. Sidhu, W. J. Miller, and S. L. O'Neill. Evidence for a global *Wolbachia* replacement in *Drosophila melanogaster*. *Current Biology*, 15:1428–1433, 2005. 9

T. Rigaud and P. Juchault. Conflict between feminizing sex ratio distorters and an autosomal masculinizing gene in the terrestrial isopod *Armadillidium vulgare* Latr. *Genetics*, 133:247–252, 1993. 13

T. Rigaud, C. Soutygrosset, R. Raimond, J. P. Mocquard, and P. Juchault. Feminizing endocytobiosis in the terrestrial crustacean *Armadillidium vulgare* Latr. (Isopoda) - recent acquisitions. *Endocytobiosis and Cell Research*, 7:259–273, 1991. 8, 13

T. Rigaud, P. Juchault, and J. Mocquard. The evolution of sex determination in isopod crustaceans. *BioEssays*, 19:409–416, 1997. 13

M. G. Riparbelli, R. Giordano, and G. Callaini. Effects of *Wolbachia* on sperm maturation and architecture in *Drosophila simulans* Riverside. *Mechanisms of Development*, 124:699–714, 2007. 16

M. G. Ritchie. Sexual selection and speciation. *Annual Review of Ecology, Evolution, and Systematics*, 38:79–102, 2007. 21

M. G. Ritchie, E. J. Halsey, and J. M. Gleason. *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou & Hall cycles in *D. melanogaster* song. *Animal Behaviour*, 58:649–657, 1999. 119

F. Rousset. Genetic differentiation in populations with different classes of individuals. *Theoretical Population Biology*, 55:297–308, 1999. 64

F. Rousset, D. Bouchon, B. Pintureau, P. Juchault, and M. Solignac. *Wolbachia* endosymbionts responsible for various alterations of sexuality in arthropods. *Proceedings of the Royal Society London B*, 250:91–98, 1992. 8

F. Rousset, H. R. Braig, and S. L. O'Neill. A stable triple *Wolbachia* infection in *Drosophila* with nearly additive incompatibility effects. *Heredity*, 82:620–627, 1999. 11

L. Sagan. On the origin of mitosing cells. *Journal of Theoretical Biology*, 14:225–274, 1967. 1

P. Schofield. Spatially explicit models of Turelli-Hoffmann *Wolbachia* invasive wave fronts. *Journal of Theoretical Biology*, 215:121–131, 2002. 28, 29, 61, 62

L. R. Serbus and W. Sullivan. A cellular basis for *Wolbachia* recruitment to the host germline. *PLoS Pathogens*, 3:e190, 2007. 11

L. R. Serbus, C. Casper-Lindley, F. Landmann, and W. Sullivan. The genetics and cell biology of *Wolbachia*-host interactions. *Annual Review of Genetics*, 42: 683–707, 2008. 11, 15, 16

M. R. Servedio. Reinforcement and the genetics of nonrandom mating. *Evolution*, 54:21–29, 2000. 21, 83, 87, 89, 92, 98, 118, 127, 151

M. R. Servedio and M. Kirkpatrick. The effects of gene flow on reinforcement. *Evolution*, 51:1764–1772, 1997. 92

M. R. Servedio and M. A. F. Noor. The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, 34:339–364, 2003. 21, 125

D. D. Shoemaker and J. Jaenike. Habitat continuity and the genetic structure of *Drosophila* populations. *Evolution*, 51:1326–1332, 1997. 138

D. D. Shoemaker, V. Katju, and J. Jaenike. *Wolbachia* and the evolution of reproductive isolation between *Drosophila recens* and *Drosophila subquinaria*. *Evolution*, 53:1157–1164, 1999. 24, 28, 61, 122, 123, 129, 139

D. D. Shoemaker, K. G. Ross, L. Keller, E. L. Vargo, and J. H. Werren. *Wolbachia* infections in native and introduced populations of fire ants (Solenopsis spp.). *Insect Molecular Biology*, 9:661–673, 2000. 83

D. D. Shoemaker, M. Ahrens, L. Sheill, M. Mescher, L. Keller, and K. G. Ross. Distribution and prevalence of *Wolbachia* infections in native populations of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae). *Environmental Entomology*, 32:1329–1336, 2003. 83

S. P. Sinkins and H. C. J. Godfray. Use of *Wolbachia* to drive nuclear transgenes through insect populations. *Proceedings of the Royal Society London B*, 271: 1421–1426, 2004. 8

S. P. Sinkins and F. Gould. Gene drive systems for insect disease vectors. *Nature Reviews Genetics*, 7:427–435, 2006. 8

S. P. Sinkins and S. L. O'Neill. *Wolbachia* as a vehicle to modify insect populations. In A. M. Handler and A. A. James, editors, *Insect transgenesis: methods and applications*, pages 271–288. CRC Press, Boca Raton, 2000. 8

S. P. Sinkins, C. F. Curtis, and S. L. O'Neill. The potential application of inherited symbiont systems to pest control. In S. L. O'Neill, A. A. Hoffmann, and J. H. Werren, editors, *Influential passengers*, pages 155–175. Oxford University Press, Oxford, 1997. 62

S. P. Sinkins, T. Walker, A. R. Lynd, A. R. Steven, B. L. Makepeace, H. C. J. Godfray, and J. Parkhill. *Wolbachia* variability and host effects on crossing type in *Culex* mosquitoes. *Nature*, 436:257–260, 2005. 12, 16, 139

H. G. Spencer, B. H. McArdle, and D. M. Lambert. A theoretical investigation of speciation by reinforcement. *American Naturalist*, 128:241–262, 1986. 83

R. Stouthamer and D. J. Kazmer. Cytogenetics of microbe-associated parthenogenesis and its consequences for gene flow in *Trichogramma* wasps. *Heredity*, 73: 317–327, 1994. 13

R. Stouthamer and J. H. Werren. Microbes associated with parthenogenesis in wasps of the genus *Trichogramma*. *Journal of Invertebrate Pathology*, 61:6–9, 1993. 13

R. Stouthamer, R. F. Luck, and W. D. Hamilton. Antibiotics cause parthenogenetic *Trichogramma* (Hymenoptera/Trichogrammatidae) to revert to sex. *Proceedings of the National Academy of Sciences of the USA*, 87:2424–2427, 1990. 8, 13

R. Stouthamer, J. A. J. Breeuwer, R. F. Luck, and J. H. Werren. Molecular identification of microorganisms associated with parthenogenesis. *Nature*, 361:66–68, 1993. 6, 8, 11

R. Stouthamer, J. A. J. Breeuwer, and G. D. D. Hurst. *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. *Annual Review of Microbiology*, 53:71–102, 1999. 10, 11, 15

Y. Tao, D. L. Hartl, and C. C. Laurie. Sex-ratio segregation distortion associated with reproductive isolation in *Drosophila*. *Proceedings of the National Academy of Sciences of the USA*, 98:13183–13188, 2001. 127

M. J. Taylor and A. Hoerauf. *Wolbachia* bacteria of filarial nematodes. *Parasitology Today*, 15:437–442, 1999. 11

P. D. Taylor. Allele-frequency change in a class-structured population. *American Naturalist*, 135:95–106, 1990. 65, 75

L. Teixeira, Álvaro Ferreira, and M. Ashburner. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biology*, 6:e2, 2008. 11

A. Telschow. Wolbachia*, Reproduktionsparasitismus und Artbildung*. PhD thesis, Humboldt University, Berlin, Germany, 2003. 86, 87

A. Telschow, P. Hammerstein, and J. H. Werren. The effect of *Wolbachia* on genetic divergence between populations: models with two-way migration. *American Naturalist*, 160:S54–S66, 2002a. 22, 61, 64, 65, 66, 69, 82, 84, 86, 127, 180

A. Telschow, P. Hammerstein, and J. H. Werren. Effects of *Wolbachia* on genetic divergence between populations: mainland-island model. *Integrative and Comparative Biology*, 42:340–351, 2002b. 61, 64, 73, 84, 86

A. Telschow, P. Hammerstein, and J. H. Werren. The effect of *Wolbachia* versus genetic incompatibilities on reinforcement and speciation. *Evolution*, 59:1607–1619, 2005a. 62, 86, 89, 92, 118, 119

A. Telschow, N. Yamamura, and J. H. Werren. Bidirectional cytoplasmic incompatibility and the stable coexistence of two *Wolbachia* strains in parapatric host populations. *Journal of Theoretical Biology*, 235:265–274, 2005b. 28, 35, 62, 86, 119

A. Telschow, J. Engelstädter, N. Yamamura, P. Hammerstein, and G. D. D. Hurst. Asymmetric gene flow and constraints on adaptation caused by sex ratio distorters. *Journal of Evolutionary Biology*, 19:869–878, 2006. 61

A. Telschow, M. Flor, Y. Kobayashi, P. Hammerstein, and J. H. Werren. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: mainland-island model. *PLoS ONE*, 2:e701, 2007. 27, 28, 61, 64, 65, 73, 85, 87, 107, 130

U. Tram and W. Sullivan. Role of delayed nuclear envelope breakdown and mitosis in *Wolbachia*-induced cytoplasmic incompatibility. *Science*, 296:1124–1126, 2002. 15, 16

J. R. True, B. S. Weir, and C. C. Laurie. Genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics*, 142:819–837, 1996. 127

M. Turelli. Evolution of incompatibility-inducing microbes and their hosts. *Evolution*, 48:1500–1513, 1994. 61

M. Turelli and A. A. Hoffmann. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. *Genetics*, 140:1319–1338, 1995. 8, 11, 129

M. Turelli and A. A. Hoffmann. Microbe-induced cytoplasmic incompatibility as a mechanism for introducing transgenes into arthropod populations. *Insect Molecular Biology*, 8:243–255, 1999. 8

M. Turelli and A. A. Hoffmann. Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature*, 353:440–442, 1991. 28, 29, 61, 83

M. Turelli, N. H. Barton, and J. A. Coyne. Theory and speciation. *Trends in Ecology and Evolution*, 16:330–343, 2001. 20, 21

F. Vala, M. Egas, J. A. J. Breeuwer, and M. W. Sabelis. *Wolbachia* affects oviposition and mating behaviour of its spider mite host. *Journal of Evolutionary Biology*, 17:692–700, 2004. 28

T. T. M. Vandekerckhove, S. Watteyne, W. Bonne, D. Vanacker, S. Devaere, B. Rumes, J. P. Maelfait, M. Gillis, J. G. Swings, H. R. Braig, and J. Mertens. Evolutionary trends in feminization and intersexuality in woodlice (Crustacea, Isopoda) infected with *Wolbachia pipientis* (α-proteobacteria). *Belgian Journal of Zoology*, 133:61–69, 2003. 13

F. Vavre, F. Fleury, D. Lepetit, P. Fouillet, and M. Bouletreau. Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. *Molecular Biology and Evolution*, 16:1711–1723, 1999. 9

Z. Veneti, M. E. Clark, S. Zabalou, T. L. Karr, C. Savakis, and K. Bourtzis. Cytoplasmic incompatibility and sperm cyst infection in different *Drosophila-Wolbachia* associations. *Genetics*, 164:545–552, 2003. 16

M. J. Wade and L. Stevens. The effect of population subdivision on the rate of spread of parasite-mediated cytoplasmic incompatibility. *Journal of Theoretical Biology*, 167:81–87, 1994. 29, 61

T. Walker, L. Klasson, M. Sebaihia, M. Sanders, N. Thomson, J. Parkhill, and S. Sinkins. Ankyrin repeat domain-encoding genes in the *w*Pip strain of *Wolbachia* from the *Culex pipiens* group. *BMC Biology*, 5:39, 2007. 12

A. R. Weeks, M. Turelli, W. R. Harcombe, K. T. Reynolds, and A. A. Hoffmann. From parasite to mutualist: rapid evolution of *Wolbachia* in natural populations of *Drosophila*. *PLoS Biology*, 5:e114, 2007. 11

J. H. Werren. Biology of *Wolbachia*. *Annual Review of Entomology*, 42:587–609, 1997. 2, 8, 9, 15, 86, 179

J. H. Werren. *Wolbachia* and speciation. In D. J. Howard and S. H. Berlocher, editors, *Endless forms: species and speciation*, pages 245–260. Oxford University Press, Oxford, 1998. 2, 23, 86

J. H. Werren and J. Jaenike. *Wolbachia* and cytoplasmic incompatibility in mycophagous *Drosophila* and their relatives. *Heredity*, 75:320–326, 1995. 122, 123

J. H. Werren, G. D. D. Hurst, W. Zhang, J. A. Breeuwer, R. Stouthamer, and M. E. N. Majerus. Rickettsial relative associated with male killing in the ladybird beetle (*Adalia bipunctata*). *Journal of Bacteriology*, 176:388–394, 1994. 8

J. H. Werren, D. Windsor, and L. R. Guo. Distribution of *Wolbachia* among neotropical arthropods. *Proceedings of the Royal Society London B*, 262:197–204, 1995a. 8

J. H. Werren, W. Zhang, and L. R. Guo. Evolution and phylogeny of *Wolbachia* - reproductive parasites of arthropods. *Proceedings of the Royal Society London B*, 261:55–63, 1995b. 9

J. H. Werren, L. Baldo, and M. E. Clark. *Wolbachia*: master manipulators of invertebrate biology. *Nature Reviews Microbiology*, 6:741–751, 2008. 4, 10, 12, 13, 14, 16

M. J. West-Eberhard. Sexual selection, social competition, and speciation. *Proceedings of the American Philosophical Society*, 123:222–234, 1979. 21

S. Wolfram. *Mathematica*. Wolfram Research, Inc., Version 7.0, Champaign, Illinois, 2008. 3, 145

S. Wright. *Evolution and the genetics of populations. Vol. 4: Variability within and among natural populations*. University of Chicago Press, Chicago, 1978. 138

S. Wright. Isolation by distance. *Genetics*, 28:114–138, 1943. 138

C. Wu and A. W. Davis. Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. *American Naturalist*, 142:187–212, 1993. 127

M. Wu, L. V. Sun, J. Vamathevan, M. Riegler, R. Deboy, J. C. Brownlie, E. A. McGraw, W. Martin, C. Esser, N. Ahmadinejad, C. Wiegand, R. Madupu, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, A. S. Durkin, J. F. Kolonay, W. C. Nelson, Y. Mohamoud, P. Lee, K. Berry, M. B. Young, T. Utterback, J. Weidman, W. C. Nierman, I. T. Paulsen, K. E. Nelson, H. Tettelin, S. L. O'Neill, and J. A. Eisen. Phylogenomics of the reproductive parasite *Wolbachia pipientis w*Mel: a streamlined genome overrun by mobile genetic elements. *PLoS Biology*, 2:e69, 2004. 4, 8, 9, 10, 11, 12

R. Yamada, K. D. Floate, M. Riegler, and S. L. O'Neill. Male development time influences the strength of *Wolbachia*-induced cytoplasmic incompatibility expression in *Drosophila melanogaster*. *Genetics*, 177:801–808, 2007. 16

J. H. Yen and A. R. Barr. The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. *Journal of Invertebrate Pathology*, 22:242–250, 1973. 6, 14

J. H. Yen and A. R. Barr. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. *Nature*, 232:657–658, 1971. 6

S. Zabalou, M. Riegler, M. Theodorakopoulou, C. Stauffer, C. Savakis, and K. Bourtzis. *Wolbachia*-induced cytoplasmic incompatibility as a means for insect pest population control. *Proceedings of the National Academy of Sciences of the USA*, 101:15042–15045, 2004. 8

E. Zchori-Fein, R. T. Roush, and M. S. Hunter. Male production induced by antibiotic treatment in *Encarsia formosa* (Hymenoptera: Aphelinidae), an asexual species. *Cellular and Molecular Life Sciences*, 48:102–105, 1992. 13

E. Zchori-Fein, S. J. Perlman, S. E. Kelly, N. Katzir, and M. S. Hunter. Characterization of a '*Bacteroidetes*' symbiont in *Encarsia* wasps (Hymenoptera: Aphelinidae): proposal of '*Candidatus* Cardinium hertigii'. *International Journal of Systematic and Evolutionary Microbiology*, 54:961–968, 2004. 12

Z. Zhang, T. M. Hambuch, and J. Parsch. Molecular evolution of sex-biased genes in *Drosophila*. *Molecular Biology and Evolution*, 21:2130–2139, 2004. 127

W. Zhou, F. Rousset, and S. L. O'Neill. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proceedings of the Royal Society London B*, 265:509–15, 1998. 9, 11

Publications

Articles in peer reviewed journals:

M. Flor, P. Hammerstein, and A. Telschow. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and the stability of infection polymorphism in parapatric host populations. *Journal of Evolutionary Biology*, 20:696–706, 2007.

A. Telschow, M. Flor, Y. Kobayashi, P. Hammerstein, and J. H. Werren. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: Mainland-island model. *PLoS ONE*, 2:e701, 2007.

J. Dannowski*, M. Flor*, A. Telschow, and P. Hammerstein. The effect of sibmating on the infection dynamics of male-killing bacteria. *Evolution*, 63:2525–2534, 2009.
* authors contributed equally

In preparation:

M. Flor, A. Telschow, and P. Hammerstein. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: Runaways and reinforcement.

M. Flor, A. Telschow, and P. Hammerstein. Modeling of the *Drosophila recens* and *D. subquinaria* speciation: Cytoplasmic incompatibility, hybrid male sterility, and sexual selection.

Poster presentations:

M. Flor, A. Telschow, Y. Kobayashi, P. Hammerstein, and J. H. Werren. Three problems of speciation via unidirectional CI ... and why it might still happen. Poster, *5th International Wolbachia Conference*, Kolymbari (Crete), Greece, 2008.

Berlin, den 27. Januar 2010

Matthias Flor

Deutsche Zusammenfassung

In dieser Arbeit wird die Rolle des intrazellulären Bakteriums *Wolbachia* in Artbildungsprozessen seiner Wirtsorganismen untersucht, in denen es zytoplasmatische Inkompatibilität verursacht.

Einleitung

Wolbachia ist eine Gattung endosymbiontischer Bakterien, die maternal über das Zytoplasma vererbt wird und insbesondere im Phylum der Arthropoden weit verbreitet ist. So schätzten kürzlich Hilgenböcker et al. (2008) in einer Metaanalyse bisheriger Studien, dass etwa zwei Drittel aller Insektenarten mit *Wolbachia* infiziert sind. Da sich die Bakterien in männlichen Wirten in einer reproduktiven Sackgasse befinden, hat die Evolution die Entstehung verschiedener Formen von Reproduktionsparasitismus begünstigt. Die häufigste Form ist die zytoplasmatische Inkompatibilität ("cytoplasmic incompatibility" – CI), eine Paarungsinkompatibilität zwischen nicht infizierten Weibchen und infizierten Männchen, die auf einem Modifikations- und Rettungsmechanismus beruht (Werren, 1997), dessen molekulare und genetische Grundlagen allerdings noch unbekannt sind[1]. Die mögliche Rolle der CI in Artbildungsprozessen der Wirtsorganismen wird bereits seit ihrer Entdeckung diskutiert (Laven, 1951, 1959). In dieser Arbeit analysieren wir häufig angeführte Kritikpunkte einer solchen Rolle mit Hilfe von mathematischen Modellen, in denen Infektionsdynamik von *Wolbachia* und Populationsgenetik der Wirte kombiniert werden.

Stabilität von Infektionspolymorphismen

Grundlage für eine Rolle *Wolbachias* in Artbildungsprozessen ist die Stabilität von Infektionspolymorphismen, also der Koexistenz infizierter und nicht infizierter Wirtspopulationen, die durch Migrationsflüsse verbunden sind. Wir zeigen in einem räumlichen Model der Infektionsdynamik, dass ein solcher stabiler Polymorphismus möglich ist, wenn die Migration unterhalb einer kritischen Migrationsrate liegt und wenn *Wolbachia* die Fekundität infizierter Weibchen verringert oder die Transmission der Bakterien nicht vollkommen ist. Unter solchen Bedingungen wirkt unidirektionale CI als postzygotische Isolationsbarriere zwischen den Populationen. Wir leiten analytische Ausdrücke für die kritische Migrationsrate in Szenarien mit unidirektionaler Migration her und zeigen anhand numerischer Simulationen, dass diese Lösungen untere Abschätzungen für den Fall bidirektionaler Migration darstellen.

Reduktion des Genflusses

Ausgehend von einem stabilen Infektionspolymorphismus beleuchten wir die Auswirkungen der unidirektionalen CI auf den Genfluss zwischen den Wirtspopulationen. Wir demonstrieren analytisch mit Hilfe eines Fitness-Graph-Ansatzes für die

[1] Die beschriebene Paarungsinkompatibilität wird als unidirektionale CI bezeichnet. Bei Infektionen mit unterschiedlichen *Wolbachia*-Stämmen ist auch eine bidirektionale Inkompatibilität möglich, die in dieser Arbeit allerdings nicht behandelt wird.

reproduktiven Werte von Wirten (Kobayashi and Telschow, 2008), dass der Genfluss asymmetrisch reduziert wird. Dies lässt sich durch effektive Migrationsraten beschreiben. Der Genfluss wird stärker eingeschränkt für die Migration von infizierten in nicht infizierte Populationen als in der umgekehrten Richtung. In größeren Populationsstrukturen führt das dazu, dass infizierte Populationen in genetische Senken verwandelt werden, wohingegen nicht infizierte Populationen zu genetischen Quellen werden. Auf diese Weise können beispielsweise Adaptationen an lokale Gegebenheiten in nicht infizierten Populationen einfacher evoluieren oder beibehalten werden als in infizierten Populationen. Unsere Formeln für die effektive Migration können als Erweiterung von früheren Resultaten betrachtet werden (Telschow et al., 2002a), wobei wir den kompletten Satz von Parametern berücksichtigen, die üblicherweise zur Beschreibung der Infektionsdynamik von *Wolbachia* verwendet werden. Außerdem stellen wir die Herleitung auf eine solidere mathematische Basis als bisher.

Verstärkung präzygotischer reproduktiver Isolation

Wir untersuchen ferner die Frage, ob die postzygotische Reproduktionsbarriere, die durch von *Wolbachia* induzierte CI zwischen Populationen mit unterschiedlichem Infektionsstatus errichtet wird, die Evolution von präzygotischer Isolation in Form von weiblicher Paarungspräferenz (sexuelle Selektion) begünstigt. Solche Prozesse werden als "reinforcement" bezeichnet (Dobzhansky, 1937). Wir verknüpfen in unserem Modell die Infektionsdynamik von *Wolbachia* mit einfacher Populationsgenetik der Wirte. Unsere wichtigsten Ergebnisse sind: (i) In einer nicht infizierten Population, die unidirektional Migrationszufluss aus einer infizierten Population erhält, kommt es sehr leicht zur Ausbreitung weiblicher Paarungspräferenz, die zu präzygotischer Isolation führt. (ii) Im Falle bidirektionaler Migration starten dagegen sich selbst verstärkende Prozesse der Ausbreitung von weiblicher Paarungspräferenz und präferiertem Merkmal ("Fisherian runaway" nach Fisher, 1930), wenn die Kosten von Paarungspräferenz klein sind. Diese "Fisherian runaway"-Prozesse führen entweder zur Fixierung der Paarungspräferenz oder des bevorzugten Merkmals in beiden Populationen. In keinem der Fälle entsteht präzygotische Isolation. (iii) Wenn die Kosten der Paarungspräferenz groß genug sind (aber nicht zu groß), dann kann durch die unidirektionale CI die Evolution präzygotischer Isolation verstärkt werden ("reinforcement"), auch wenn Migration in beiden Richtungen stattfindet. (iv) "Reinforcement" reduziert den Genfluss bidirektional, die Auswirkungen sind allerdings deutlich stärker für die Migration aus infizierten in nicht infizierte Populationen. Die relative Bedeutung der zytoplasmatischen Inkompatibilität zur reproduktiven Isolation nimmt dadurch ab. (v) Die Ausbreitung weiblicher Paarungspräferenz in den untersuchten Szenarien ist ein langsamer Prozess.

Anwendung des Modelles: *Drosophila* in Nordamerika

Schließlich wenden wir unsere Modelle auf den tatsächlichen Artbildungsprozess der Geschwisterarten *Drosophila subquinaria* und *D. recens* in Nordamerika an, für den in Paarungsexperimenten Muster von sexueller Selektion nachgewiesen wurden, die den erwarteten Auswirkungen von *Wolbachia* entsprechen (Jaenike et al., 2006). Wir passen die Modelle an die geographischen Gegebenheiten an und schätzen Pa-

rameterwerte ab. Wir zeigen, dass die Sterilität von Hybridmännchen kaum einen Einfluss auf die Stabilität des Infektionspolymorphismus' hat. Unsere Simulationen bestätigen, dass *Wolbachia* den gefundenen Paarungspräferenzmustern unter den gegebenen Bedingungen zu Grunde liegen kann, allerdings nur unter der Annahme von stark reduzierter Rekombination zwischen Präferenz- und Merkmalslokus. Insgesamt lassen unsere Ergebnisse vermuten, dass für die Artbildung wichtige Eigenschaften dieses *Drosophila*-Systems bisher übersehen worden sind.

Konklusion

Zusammenfassend implizieren unsere Ergebnisse, dass *Wolbachia* häufig mit der Entstehung neuer Wirtsarten verknüpft sein kann, allerdings in den meisten Fällen nur, indem die von den Bakterien bewirkte zytoplasmatische Inkompatibilität als einer von mehreren Faktoren zur reproduktiven Isolation beiträgt. Eine Verstärkung sexueller Isolation wird nur unter speziellen Bedingungen bewirkt.

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die Dissertation selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Ich habe mich nicht anderwärts um einen Doktorgrad beworben und besitze derzeit keinen Doktorgrad.

Ich bin in Kenntnis der zugrundeliegenden Promotionsordnung.

Berlin, den 27. Januar 2010

Matthias Flor